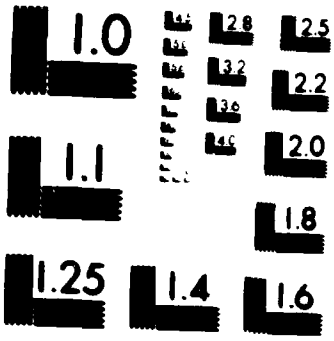
AD-A188 885 COMPUTATION OF LMS (LEAST-MEAN-SQUARE) AND MATCHED DIGITAL FILTERS FOR OPTICAL CLUTTER SUPPRESSION(U) NAVAL RESEARCH LAB WASHINGTON DC M S LONGMIRE ET AL. 1/1

UNCLASSIFIED 31 DEC 87 NRL-MR-6125 F/G 9/5 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

Naval Research Laboratory
Washington, DC 20375-5000

DTIC FILE COPY

NRL Memorandum Report 6125

AD-A188 885

Computation of LMS and Matched Digital Filters for Optical Clutter Suppression

M. S. LONGMIRE

Department of Physics and Astronomy
Western Kentucky University
Bowling Green, KY 42101

E. H. TAKKEN

Electro-optical Technology
Optical Sciences Division

December 31, 1987

DTIC
ELECTE
FEB 1 0 1988
S D
E

Approved for public release, distribution unlimited.

8 2 0 51

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release; distribution unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S) NRL Memorandum Report 6125	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Research Laboratory — 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION
6c ADDRESS (City, State, and ZIP Code) Washington, DC 20375-5000	7b ADDRESS (City, State, and ZIP Code)
8a NAME OF FUNDING / SPONSORING ORGANIZATION Naval Air Systems Command — 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code) Washington, DC 20361-8030	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 63206N; PROJECT NO:; TASK NO:; WORK UNIT ACCESSION NO: DN330-274

11 TITLE (Include Security Classification)
Computation of LMS and Matched Digital Filters for Optical Clutter Suppression

12 PERSONAL AUTHOR(S)
Longmire, M. S.* and Takken, E. H.

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM 3/1/85 TO 3/1/86	1987 December 31	50

16 SUPPLEMENTARY NOTATION
*Western Kentucky University, Bowling Green, KY 42101

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Digital filters; Electro-optics; Optical clutter suppression; Infrared technology

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Methods of computing impulse-response weights of one- and two-dimensional matched filters and least-mean-square (LMS) filters for suppressing clutter in an electro-optic sensor's output are developed and illustrated with examples. The methods are applicable to signals from scanning or staring sensors viewing finite or point sources against variable backgrounds, provided signal shape and orientation are known. The matched-filter design technique is based on isotropic power-spectral clutter models whose parameters also must be known. Images of sensor output are assumed to provide the requisite information about signals and backgrounds. The LMS design technique is based on deterministic polynomial clutter models. An LMS filter estimates signal amplitude and, implicitly, local clutter parameters by performing a least-squares fit of a signal-plus-clutter model to the sensor output at every point of the scene. Thus clutter parameters need not be knows for LMS design, though qualitative knowledge of the background may facilitate choice of the clutter model.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL E. H. Takken	22b TELEPHONE (Include Area Code) 202-767-3153 — 22c OFFICE SYMBOL Code 6550

DD Form 1473, JUN 86 — Previous editions are obsolete — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

CONTENTS

I. INTRODUCTION 1

II. 1-D AND 2-D MATCHED FILTERS 3

A. Analysis and Example 3
B. Discussion 6

III. 1-D LMS FILTERS 7

A. Basic Analysis and Example 8
B. Further Analysis and Discussion 10

IV. 2-D LMS FILTERS 14

A. Basic Analysis and Example 15
B. Further Analysis and Discussion 19

V. SUMMARY 21

TABLES 24-28

FIGURES 29-30

APPENDIX A — Clutter Models 31

APPENDIX B — Convolution Arrays and Co-ordinates 34

APPENDIX C — Approximation of Derivatives and Derivative Operators 37

REFERENCES 44

QUALITY INSPECTED 2

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

COMPUTATION OF LMS AND MATCHED DIGITAL FILTERS FOR OPTICAL CLUTTER SUPPRESSION

I. INTRODUCTION

Matched spatial filters are commonly used for suppressing clutter and detecting signals in either real-time or imaged output of electro-optic sensors. Least-mean-square (LMS) spatial filters have been so-used in infrared (IR) applications. This article extends certain known methods of calculating digital impulse responses for these filters so that the methods are applicable with both one- and two-dimensional (1-D and 2-D) forms of rather general signal and clutter models. Concomitantly, 2-D LMS filters are derived for the first time, and their make-up and operation are investigated and explained. A primary objective of this article is to enable readers to compute digital impulse-response weights of an LMS or matched filter for any signal and clutter representable by models of the assumed types. Accordingly, the design methods are described fully and illustrated with examples so that they can be applied easily in diverse conditions.

The methods can be used with signals from scanning or staring sensors viewing extended or point sources against variable backgrounds, but signal shape and orientation must always be known. Since both point and extended sources have finite images, a signal from either can be resolved (consist of more than one nonzero sample). The signal shape then varies depending on where the image falls in relation to the sample sites. In addition, an extended source eclipses the background, which results in variable source contrast and further signal variability. Thus the design methods are useful if signal shape is representable by an average or if detection of a particular signal is imperative. If the signal is extremely variable and must be detected in all its guises, the present methods fail. For reasons to be explained, the matched-filter design technique requires knowledge of clutter parameters, whereas the LMS design technique does not, though qualitative knowledge of the background may be useful for choosing the LMS clutter model. Needed information about signal and clutter is assumed to be available from images of sensor output, obtained either experimentally or theoretically.

Matched filters are far older and better known than LMS filters, having been invented by D. O. North in 1943.[1] The first matched filter was designed for detection of radar signals in white noise. Subsequently the matched-filter concept was extended to other types of noise and to optical signals.[2–5] The first matched filters were for continuous input and were difficult to make. Development of surface-acoustic-wave delay lines and digital electronics overcame the manufacturing problems and greatly increased the use of matched filters for processing both continuous and sampled input.[6] Today, 2-D digital matched filters are widely used for pattern recognition and image processing.[7,8]

Matched filters result from maximizing the ratio of filtered signal energy to average filtered noise power. There are various ways to derive digital filter weights from this criterion. The method elaborated here was introduced for IR signals by Otazo, Parenti, and Tung, who employed an anisotropic clutter model based on analysis of measured IR backgrounds.[9,10] Their work resulted in the now well known *fourth-derivative filters*, obtained by estimating the mixed partial derivative

Manuscript approved August 3, 1987.

$(\partial^4 s / \partial x^2 \partial y^2)$ at sampled points of the signal s. Kay and others, who also analyzed measured backgrounds, have suggested that in some respects an isotropic clutter model may be preferable.[11-13] In this article, the design technique of Otazo et al. is extended by using a generalized form of Kay's isotropic model. The results are presented as operator equations for the filter weights and a table of operator arrays.

Recently W. N. Peters derived theoretical expressions for matched-filter transfer functions by maximizing the signal-to-noise ratio (SNR) of the entire electro-optical system from the background and source through the digital filter.[14] The 2-D filters obtained in this way are called *type A* or *type B* depending on whether signal and clutter are assumed to be known in the object plane or at the sensor's output; 1-D filters designed from sensor output are called *type C*. The matched filters of the present article are Peters' type B and type C since they are based on sensor output.

LMS filters are obtained from least-squares fits of signal-plus-clutter models to sensor output. In fact, an LMS filter implements a regression analysis of the output and thereby estimates the amount of model signal present. Takken et al. introduced the 1-D LMS filter for continuous signals and extended the continuous results to sampled data.[15] Here analysis of the discrete case is simplified by applying the least-squares method directly to the sampled data. Equations are derived for the weights of 1-D LMS filters based on first- through fifth-degree clutter polynomials. The design technique for discrete 2-D LMS filters is then worked out. It follows the same pattern as the 1-D technique, but there are more equations with more terms. Equations are given for the weights of 2-D LMS filters based on linear through cubic clutter models. Equations for higher-order 2-D filters are easily derived but are not given because they are too lengthy.

Frequency-domain characteristics of LMS filters receive considerable attention in reference 15. There in Appendix A it is shown that 1-D LMS filters are fixed bandpass filters having a special property—the ability to block certain waveforms whose complexity increases with the degree of the clutter polynomial. Frequency-domain characteristics of LMS filters are not further investigated in this article because they are not needed for filter design and are adequately treated in reference 15. Moreover, emphasizing the space-domain analysis calls attention to properties of LMS filters that are not apparent in the frequency domain and are not mentioned in reference 15.

Because of a similarity in names, it is advisable to compare and contrast LMS filters with *LMS adaptive filters*. The latter devices employ an *LMS algorithm* that adjusts their weights to minimize the difference between the filter output and an externally supplied desired response.[16] In the present context, the signal shape and orientation are the desired response The LMS algorithm does not use the method of least squares directly but rather approximates the solution of the Wiener-Hopf least-squares equations, hence the names of the algorithm and filter.[17] The LMS algorithm is known to converge with uncorrelated and correlated stationary input, and with some correlated nonstationary input, but it is not known to converge unconditionally.[17,18] Since the algorithm acts iteratively and in effect employs statistical samples of limited size, the filter exhibits *settling time* and *misadjustment* effects.[16,17] These could lead to problems with rapidly varying nonstationary backgrounds. The LMS filter discussed here is a much simpler device whose weights are fixed. It nonetheless conforms to the input because the parameters of its signal-plus-clutter model vary spontaneously to fit the background and signal, if any, within the filter. The quality of the fit depends on the model's fidelity to the local background and so varies over the scene. Since the LMS filter solves least-squares regression equations exactly, there are no questions of convergence, settling time, and misadjustment. These features make the LMS filter well suited for suppression of optical clutter, which can vary rapidly in almost any fashion.

The best known device derived by a least-squares analysis is the Wiener filter. Its weights are the exact solution of the Wiener-Hopf equations—the solution approximated by the LMS algorithm. Wiener filters and LMS filters differ in the information needed to determine the filter weights and in

the type of output. To determine the weights of a Wiener filter, it is necessary to know the autocorrelation of the input and its cross correlation with the signal.[19] To determine LMS-filter weights, only signal and clutter models are required. No statistical properties of the input are needed. The output of a Wiener filter estimates the signal itself. The output of an LMS filter estimates a signal parameter, the amplitude.

Finally, a few words about the organization of this article will be helpful. Matched filters, 1-D LMS filters, and 2-D LMS filters are discussed separately in sections (II, III, IV), each having subsections A and B. Subsections A give the basic analyses required for calculating filter weights, with numerical examples that can be worked through to test understanding of the analyses. Subsections B discuss the equations and filters and analyze LMS filters for symmetric signals, which present special problems. Those who only want to calculate filter weights can omit subsections B unless they have trouble with their computations. In that case they may wish to read these subsections and will be in a better position to do so. Three appendices give detailed information about matched-filter clutter models, convolution arrays, and construction of these arrays to approximate derivative operators. Those who wish to calculate filter weights for conditions different from the ones assumed in sections II-IV may need to consult the appendices.

II. 1-D AND 2-D MATCHED FILTERS

Matched filters can be derived for sampled or continuous sensor output by using either clutter data from a single background or a clutter model that approximates data from a large set of backgrounds.[9,10] The data-based and model-based approaches are complementary. A matched filter based on data from a single background has the greatest signal-to-clutter ratio (SCR) achievable with a linear filter, but only for the data employed. A matched filter based on a clutter model is more widely applicable; it has nearly maximum SCR against all backgrounds reasonably well represented by the model. The performance of a model-based filter against a specified background can be evaluated by comparison with the data-based filter derived specifically for that background. In this way, other investigators have found that matched filters based on Eq. (A6) of Appendix A have SCRs not more than 4 dB below the SCRs of the best possible matched filters against those backgrounds.[9] Consequently, the matched filters developed in the present work are based on power-spectral clutter models somewhat more general than Eq. (A6). Continuous filters are derived first and then converted to digital form. The impulse responses rather than the transfer functions are digitized because the filters of interest are small—they act on no more than about 50 data at once. In such cases the filters are more efficiently implemented in the space domain than in the frequency domain (ref. 8, sec. 11.3).

(A) Analysis and Example

The one- and two-dimensional filters will be developed jointly so that the simpler 1-D analysis will clarify the more intricate 2-D treatment. The first step is to derive the filter transfer functions by maximizing the signal-to-average-clutter ratios. The derivations are not given here because they can be found in many readily available sources, *e.g.*, refs. 7,8,20-22. For present purposes the essential results are the equations of the transfer functions and impulse responses together with definitions of the quantities involved. The 1-D and 2-D transfer functions are

$$H(\omega) = c\frac{S^*(\omega)}{N(\omega^2)}e^{-j\omega v} \quad \text{and} \quad H(u, v) = k\frac{S^*(u, v)}{N(\Omega^2)}e^{-j(ux+vy)}, \tag{1a,b}$$

with impulse responses

$$h(\beta) = F^{-1}[H(\omega)] = \frac{c}{2\pi}\int_{-\infty}^{\infty} \frac{S^*(\omega)}{N(\omega^2)}\, e^{-j\omega y} e^{j\omega\beta} d\omega, \text{ and} \tag{2a}$$

$$h(\alpha, \beta) = F^{-1}[H(u, \mathrm{v})]$$

$$= \frac{k}{4\pi^2}\iint_{-\infty}^{\infty} \frac{S^*(u, \mathrm{v})}{N(\Omega^2)}\, e^{-j(ux+\mathrm{v}y)}\, e^{j(u\alpha+\mathrm{v}\beta)} du d\mathrm{v}. \tag{2b}$$

In these equations

c, k	=	arbitrary constants;
j	=	imaginary unit, $j^2 = -1$;
x, y	=	co-ordinates of maximum signal and maximum signal-to-average-clutter ratio;
ω	=	1-D spatial frequency (radians/unit length);
u, v	=	cartesian components of 2-D spatial frequency $\hat{\Omega} = \hat{u} + \hat{\mathrm{v}}$;
$S^*(\omega), S^*(u, \mathrm{v})$	=	complex conjugates of signal spectra at the sensor output;
$N(\omega^2), N(\Omega^2)$	=	clutter power-spectral densities at the sensor output;
F, F^{-1}	=	Fourier transformation;
α, β	=	Fourier-transform parameters and convolution variables corresponding to x, y.

The following relations with $\cdot$ denoting the vector inner product further define some of the quantities.

$$\Omega^2 = \hat{\Omega}\cdot\hat{\Omega} = u^2 + \mathrm{v}^2 \qquad \hat{\Omega}\cdot(\hat{x} + \hat{y}) = ux + \mathrm{v}y \qquad \hat{\Omega}\cdot(\hat{\alpha} + \hat{\beta}) = u\alpha + \mathrm{v}\beta \tag{3a,b,c}$$

Explicit forms of the power-spectral densities are needed to evaluate the transforms for the impulse responses. The power spectra employed are

$$N(\omega^2) = N_0/[1 + aL^2\omega^2 + bL^4\omega^4 + cL^6\omega^6] \text{ and} \tag{4a}$$

$$N(\Omega^2) = N_0/[1 + aL^2\Omega^2 + bL^4\Omega^4 + cL^6\Omega^6], \tag{4b}$$

where L is the clutter correlation length. These equations are spatially isotropic clutter models. Their source and the reasons for their choice are discussed in Appendix A. Substituting Eqs. (4a,b) and (3a) into Eqs. (2a,b) gives

$$h(\beta) = cN_0F^{-1}\{[1 + aL^2\omega^2 + bL^4\omega^4 + cL^6\omega^6]S^*(\omega)e^{-j\omega y}\}, \text{ and} \tag{5a}$$

$$h(\alpha, \beta) = kN_0F^{-1}\{[1 + aL^2(u^2 + \mathrm{v}^2) + bL^4(u^4 + 2u^2\mathrm{v}^2 + \mathrm{v}^4)$$

$$+ cL^6(u^6 + 3u^4\mathrm{v}^2 + 3u^2\mathrm{v}^4 + \mathrm{v}^6)]S^*(u, \mathrm{v})e^{-j(ux+\mathrm{v}y)}\}. \tag{5b}$$

The first terms in Eqs. (5a,b) are the same as the impulse responses of matched filters in white noise.

$$F^{-1}\{S^*(\omega)e^{-j\omega y}\} = s(y - \beta) \text{ and} \tag{6a}$$

$$F^{-1}\{S^*(u, \mathrm{v})e^{-j(ux+\mathrm{v}y)}\} = s(x - \alpha, y - \beta). \tag{6b}$$

In each case s is the known signal. If Eqs. (6a,b) are differentiated repeatedly with respect to x and y, the even derivatives can be identified with the remaining terms of Eqs. (5a,b). The resulting impulse responses with inconsequential parameters omitted are

$$h(\beta) = \left[1 - aL^2\frac{d^2}{dy^2} + bL^4\frac{d^4}{dy^4} - cL^6\frac{d^6}{dy^6}\right] s(y - \beta), \tag{7a}$$

$$h(\alpha, \beta) = \left[1 - aL^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + bL^4\left(\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2 \partial y^2} + \frac{\partial^4}{\partial y^4}\right)\right.$$

$$\left. -cL^6\left(\frac{\partial^6}{\partial x^6} + 3\frac{\partial^6}{\partial x^4 \partial y^2} + 3\frac{\partial^6}{\partial x^2 \partial y^4} + \frac{\partial^6}{\partial y^6}\right)\right] s(x - \alpha, y - \beta) \tag{7b}$$

$$= [1 - aL^2\nabla^2 + bL^4\nabla^4 - cL^6\nabla^6]s(x - \alpha, y - \beta), \tag{7c}$$

where ∇^2 represents the Laplacian operator $(\partial^2/\partial x^2) + (\partial^2/\partial y^2)$.

These analog impulse responses can be converted to digital form by approximating the derivatives numerically at sampled points of the signals. Appendices B and C discuss the approximations and their representation by convolution arrays $[\omega^2]$, $[\nabla^2]$, etc. The digital matched-filter impulse responses written in terms of these arrays are

$$[h(-\beta)] = \{1 - aL^2[\omega^2] + bL^4[\omega^4] - cL^6[\omega^6]\} * [s(\beta)] \quad \text{and} \tag{8a}$$

$$[h(-\alpha, -\beta)] = \{1 - aL^2[\nabla^2] + bL^4[\nabla^4] - cL^6[\nabla^6]\} * [s(\alpha, \beta)], \tag{8b}$$

where the asterisk denotes convolution. $[s(\beta)]$ and $[s(\alpha, \beta)]$ are arrays of signal samples. With tabulated signals, the impulse responses need not be delayed to satisfy causality; hence x and y of Eqs. (7a-c) have been equated to zero. In keeping with Eqs. (7a-c), the notation indicates that the operators act on unreversed signals, and the resulting arrays are rotated 180° about their centers to obtain the impulse responses. The impulse-response arrays are rotated another 180° before being applied to filter input, as explained in Appendix B. Table 1 gives array approximations of the derivative operators for equal sampling intervals in the x and y directions. Appendix C explains how to construct the operator arrays when the x and y sampling intervals are not equal.

An example will illustrate the use of Eq. (8b) to determine an impulse response. To simplify the example, it will be assumed that the second-derivative term dominates Eq. (8b), allowing the other terms to be ignored. Let the signal with unit total strength be

$$[s(\alpha, \beta)] = \frac{1}{6}\begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}. \tag{9}$$

This could be a fixed signal or the average of a variable signal. Since the signal is known to be zero beyond the listed values, peripheral zeros can be added to permit convolution with the Laplacian operator. Equal x and y sampling intervals are assumed in order to allow use of $[\nabla^2]$ from Table 1. Then from Eq. (8b) and Appendix B

$$-\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} 0_{11} & \cdot & \cdot & \cdot \\ \cdot & 0 & 1 & \cdot \\ \cdot & 3 & 2 & \cdot \\ \cdot & \cdot & \cdot & 0_{66} \end{bmatrix} = -\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 4 & -2 & 1 \\ 3 & -10 & -4 & 2 \\ 0 & 3 & 2 & 0 \end{bmatrix}. \tag{10}$$

Reversing the resulting array gives the unnormalized impulse response,

$$[h] = -\begin{bmatrix} 0 & 2 & 3 & 0 \\ 2 & -4 & -10 & 3 \\ 1 & -2 & 4 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{11}$$

The signal output from this unnormalized filter is

$$-\begin{bmatrix} 0 & 2 & 3 & 0 \\ 2 & -4 & -10 & 3 \\ 1 & -2 & 4 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} * \frac{1}{6}\begin{bmatrix} 0_{11} & \cdot & \cdot & \cdot \\ \cdot & 0 & 1 & \cdot \\ \cdot & 3 & 2 & \cdot \\ \cdot & \cdot & \cdot & 0_{88} \end{bmatrix} = -\frac{1}{6}\begin{bmatrix} 0 & 0 & 2 & 3 & 0 \\ 0 & 8 & 9 & -4 & 3 \\ 6 & -7 & -40 & -7 & 6 \\ 3 & -4 & 9 & 8 & 0 \\ 0 & 3 & 2 & 0 & 0 \end{bmatrix}, \tag{12}$$

with maximum strength (40/6). The normalized impulse response that passes the signal with maximum strength equal to the input total strength is

$$[h] = -\frac{3}{20}\begin{bmatrix} 0 & 2 & 3 & 0 \\ 2 & -4 & -10 & 3 \\ 1 & -2 & 4 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{13}$$

The normalization, or gain, of (3/20) is arbitrary. Since gain does not affect signal-to-clutter ratio, any gain is permissible. Consequently, the most convenient normalization criterion can be used in each case.

For simplicity only, one term of the impulse response was kept in the example. In practice the terms kept are determined primarily by the relative values of the power-spectral coefficients (1, aL^2, bL^4, cL^6). These are best obtained by fitting the model Eqs. (4a,b) to measured power spectra of the backgrounds involved. Unless one of the coefficients is much larger than the others, the impulse response is a sum of two or more terms properly weighted by the coefficients. However, the unity term is usually negligible and can be omitted.[9,10] Furthermore, unless it is omitted, the impulse-response weights do not sum to zero and, therefore, the filter output does not have zero mean locally and globally. This property is essential for correct operation of the usual type of adaptive-threshold sensor that follows the filter.[23,24] Omitting the unity term from the impulse responses corresponds in the frequency domain to omitting the same term from the power-spectral models, and this gives transfer functions Eqs. (1a,b) that are zero at zero frequency.

(B) Discussion

The unity term is sometimes called the white-noise term because it has the same form as the impulse response of a matched filter in white noise. Nonetheless, this term comes from the constant of the clutter model, and that constant represents, not white noise, but power-spectral density of pure clutter at zero frequency. The filters of Eqs. (8a,b) are derived from models that take no account of white noise. Yet the clutter waveform actually is the sum of a pure-clutter component from the variable background radiance and a white-noise component from photon arrival fluctuations and charge fluctuations in the detector and electronics. The true clutter power spectrum is, therefore, the sum of a pure-clutter term, a white-noise term, and cross-spectral terms between the pure clutter and the white noise.[25,26] The pure-clutter term is much larger than the white-noise and cross-spectral terms except at high frequencies. Can the small terms be neglected? Figure 4 of reference 15 displays SNR as a function of noise type for a matched filter in white noise and a matched filter in $(1/f^2)$

clutter. The SNRs of the two filters differ by 4.5 dB when both operate in white noise. Thus, including the small terms in the $(1/f^2)$ filter's design could not improve its white-noise performance more than 4.5 dB, and probably would improve it much less. If this example is typical, and the indicated deficit in white-noise performance in acceptable, then Eqs. (4a,b) are adequate for deriving clutter-suppression filters.

It is also worthwhile to consider the relation between the clutter models and the impulse-response terms whose coefficients are aL^2, bL^4, cL^6. Comparison of Eqs. (4), (5), and (7) shows that ω^k or $u^m v^n$ in the models leads to $j^k(d^k/dy^k)$ or $j^{m+n}(\partial^{m+n}/\partial x^m \partial y^n)$ in the impulse responses for even k, m, and n. Because of that fact, the impulse-response terms with coefficients aL^2, bL^4, cL^6 are called, respectively, matched filters for $(1/f^2)$, $(1/f^4)$, and $(1/f^6)$ clutter. Odd powers of frequency have been left out of the clutter models because they do not give such simple, general, and easily evaluated terms in the impulse response (Appendix A).

The isotropy of the models also affects the impulse-response terms. Power spectra of isotropic models contain the spatial frequencies u, v only in the combination $\Omega^2 = u^2 + v^2$ (ref. 27, sec. 10.3.2; ref. 28, pp. 244-248). For comparison, consider the anisotropic power spectrum,

$$N(u^2, v^2) = N_0/[(1 + L_x^2 u^2)(1 + L_y^2 v^2)], \tag{14}$$

which leads to the matched-filter impulse response,

$$h(\alpha, \beta) = \left[1 - L_x^2 \frac{\partial^2}{\partial x^2} - L_y^2 \frac{\partial^2}{\partial y^2} + L_x^2 L_y^2 \frac{\partial^4}{\partial x^2 \partial y^2}\right] s(x - \alpha, y - \beta). \tag{15}$$

The last term on the right side usually predominates; in that case numerical approximation of $(\partial^4/\partial x^2 \partial y^2)$ by $[\omega^{22}]$ of Table 1 gives the fourth-derivative filters of Otazo et al.[9,10] Equation (15), unlike Eq. (7b), contains no sixth derivatives and no homogeneous partial derivatives above the second. Both differences are easily identified with the powers of u and v in the respective clutter models. Another notable difference is that the anisotropic impulse response does not contain the Laplacian operator, whereas the isotropic response Eq. (7c) can be written in powers of that operator. This is entirely due to the latter clutter model's isotropy and the absence of odd powers of frequency, specifically to the facts that the correlation length is a single constant, and the spatial frequency components occur only in the combination $(u^2 + v^2)^p$, $p = 1, 2, 3$.

In what circumstances might one prefer an isotropic filter or vice versa? References 9-13 discuss this and related questions, so comments here will be brief. If an optical sensor's output has a predominant anisotropy, a filter mated to that anisotropy is preferable. A sensor could produce such an output if it views only a small class of anisotropic scenes, or if its construction or perspective induce a strong anisotropy in the output. On the other hand, a sensor may view a large class of scenes with different equally probable anisotropies, and it may be constructed and operated so as to induce no particular anisotropy in the output. In that case the output is likely to be isotropic on average, making an isotropic filter desirable. Stated somewhat differently, when the anisotropy of the sensor output is unknown and/or highly variable, an isotropic matched filter could be a desirable compromise on average. However, an LMS filter may be preferable in these circumstances since its clutter-model parameters change to suit the local background.

III. 1-D LMS FILTERS

Although 2-D filters probably are used more often than 1-D filters for electro-optical signal processing, 1-D LMS design is studied at length in this section because its key features carry over to and illuminate the more intricate 2-D problem, which is not easily treated in the same detail. Subsection III A develops the 1-D design procedure and gives all that is needed to calculate an LMS digital filter for any sampled 1-D signal. Subsection III B elaborates certain aspects of the 1-D analysis and filters.

(A) Basic Analysis and Example

In the LMS design procedure a deterministic model of signal-plus-clutter is fit by least squares to the observed sensor output. The signal is represented by the product of an amplitude A and a shape function s. The clutter is assumed to be a continuous function of position y and to be representable within the filter's limited span by a few terms of the function's Maclaurin-series expansion. With K samples per signal length and a clutter polynomial of degree M, the signal-plus-clutter model can be written

$$f_j = As_j + \sum_{m=0}^{M} B^m y_j^m,\ j = 1, 2, 3 \ldots K, \text{ and } K \geq (M + 2). \tag{16}$$

The superscript of B^m is not an exponent; it simply associates B^m with a power of y_j in the clutter model. Equation (16) is fit to the K observations within the filter by adjusting the parameters A and B^m. Consequently, the signal amplitude and clutter coefficients vary over the scene, whereas the signal shape and orientation are fixed. In this way the signal-plus-clutter model adjusts to the input, as illustrated in Fig. 1 which shows how linear, quadratic, and cubic clutter polynomials might fit the background at three filter positions.

At any point within the filter, the difference between the signal-plus-clutter model and the observed sensor output v_j is

$$e_j = f_j - v_j,\ j = 1, 2, 3, \ldots K. \tag{17}$$

Equations (17) are called equations of condition. If the parameters and observations are equal in number, these equations can be solved for parameters that make all differences zero. Usually, however, least-squares equations are derived by minimizing the sum of the squared differences with respect to the parameters, and then the least-squares equations are solved for the best-fit parameters. Either way, the solution for the signal amplitude A provides the relative weights of an LMS digital filter whose output estimates the amplitude of the model signal in the K samples contained by the filter.

Least-squares equations will now be derived for a quintic signal-plus-clutter model. To simplify notation the subscript k will imply summation over the observations, e.g.,

$$e_k^2 \equiv \sum_{j=1}^{K} (f_j - v_j)^2. \tag{18}$$

Differentiating e_k^2 with respect to the adjustable parameters and equating the derivatives to zero, one obtains

$$\sum_{j=1}^{K} e_j \frac{\partial e_j}{\partial A} \equiv e_k \frac{\partial e_k}{\partial A} = 0 \text{ and } e_k \frac{\partial e_k}{\partial (B^m)} = 0,\ m = 0, 1, 2, 3, 4, 5. \tag{19}$$

Combining Eqs. (16,17,19) gives the desired least squares equations, which in matrix form are

$$\begin{vmatrix} s_k^2 & s_k & s_k y_k & s_k y_k^2 & s_k y_k^3 & s_k y_k^4 & s_k y_k^5 \\ s_k & K & y_k & y_k^2 & y_k^3 & y_k^4 & y_k^5 \\ s_k y_k & y_k & y_k^2 & y_k^3 & y_k^4 & y_k^5 & y_k^6 \\ s_k y_k^2 & y_k^2 & y_k^3 & y_k^4 & y_k^5 & y_k^6 & y_k^7 \\ s_k y_k^3 & y_k^3 & y_k^4 & y_k^5 & y_k^6 & y_k^7 & y_k^8 \\ s_k y_k^4 & y_k^4 & y_k^5 & y_k^6 & y_k^7 & y_k^8 & y_k^9 \\ s_k y_k^5 & y_k^5 & y_k^6 & y_k^7 & y_k^8 & y_k^9 & y_k^{10} \end{vmatrix} \begin{vmatrix} A \\ B^0 \\ B^1 \\ B^2 \\ B^3 \\ B^4 \\ B^5 \end{vmatrix} = \begin{vmatrix} s_k v_k \\ v_k \\ y_k v_k \\ y_k^2 v_k \\ y_k^3 v_k \\ y_k^4 v_k \\ y_k^5 v_k \end{vmatrix} \tag{20}$$

where, e.g., $s_k y_k^2 = \Sigma\, s_j y_j^2$ and the dummy index of summation on the right side has been changed from k to λ for clarity in later work. Equations for lower-degree signal-plus-clutter models are obtained simply by eliminating appropriate elements from Eq. (20). For example, the equation for a quadratic model is obtained by eliminating the last three rows and columns of the coefficient matrix and the last three elements of the parameter and observation vectors. The coefficients of the parameters and observations are obtained from the known values of the model signal s_j and the sample co-ordinates y_j. Both the coefficient matrix and computation of its elements are simplified by choosing the co-ordinate origin judiciously. A 1-D signal has either an odd or even number of samples which lie equally on each side of a central point. If the origin is placed at the central point, the number of co-ordinate values y_j is halved except for sign, and all matrix elements y_k^p with p odd are zero. After the coefficients have been determined, the set of linear equations is solved numerically for the best-fit signal-amplitude parameter A, and the terms of the solution are written as a convolution sum (defined and illustrated in Appendix B). Finally, the LMS filter weights are obtained by inspection of the convolution sum.

An LMS filter for the signal [5 2 1 3 4 3] determined by six throws of a fair die will be computed as an example. The signal is chosen to emphasize the fact that the procedure is applicable to any signal whatsoever. Assume that the clutter within the signal length is adequately fit by a quadratic function. The co-ordinate origin is placed at the signal center, and y_j is measured in units of the sampling interval. Thus we have:

Model	:	$f_j = As_j + (B^0 + B^1 y_j + B^2 y_j^2),\ j = 1, 2, \ldots 6$					
Indices (j)	:	1	2	3	4	5	6
Co-ordinates (y_j)	:	$-\frac{5}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$
Signal values (s_j)	:	5	2	1	3	4	3
Observations (v_j)	:	v_1	v_2	v_3	v_4	v_5	v_6

Elements are eliminated from Eq. (20) to suit the model, and the co-ordinates and signal values are substituted into the reduced equations. This gives the following set of equations for the best-fit model parameters:

$$\begin{pmatrix} 64 & 18 & -1 & 64.5 \\ 18 & 6 & 0 & 17.5 \\ -1 & 0 & 17.5 & 0 \\ 64.5 & 17.5 & 0 & 88.375 \end{pmatrix} \begin{pmatrix} A \\ B^0 \\ B^1 \\ B^2 \end{pmatrix} = \begin{pmatrix} s_\lambda v_\lambda \\ v_\lambda \\ y_\lambda v_\lambda \\ y_\lambda^2 v_\lambda \end{pmatrix}. \tag{21}$$

The solution of interest is that for the signal amplitude,

$$A = (3920 s_\lambda v_\lambda - 8085 v_\lambda + 224 y_\lambda v_\lambda - 1260 y_\lambda^2 v_\lambda)/23856. \tag{22}$$

The denominator is the determinant of the coefficient matrix; it is absorbed eventually in the filter normalization (gain). The unsigned numbers in the numerator are the minors of the elements in the first column of the determinant array. Since the denominator is irrelevant and the subscripts imply summation,

$$A = \sum_{j=1}^{6} (3920 s_j - 8085 + 224 y_j - 1260 y_j^2) v_j = \sum_{j=1}^{6} w_j v_j. \tag{23}$$

The right side of Eq. (23) can be interpreted as a convolution sum in which the parenthetical subsums are the LMS filter weights. In terms of convolution arrays [Appendix B, Eqs. (B1-B4)] the last equality of Eq. (23) is

$$A = [w_1 w_2 w_3 w_4 w_5 w_6] * [v_1 v_2 v_3 v_4 v_5 v_6] = w_6 v_1 + w_5 v_2 + w_4 v_3 + w_3 v_4 + w_2 v_5 + w_1 v_6. \tag{24}$$

By comparison of Eqs. (23) and (24),

$$w_6 = 3920 s_1 - 8085 + 224 y_1 - 1260 y_1^2 = 3080. \tag{25}$$

In the same way one finds for the other weights $w_5 = -3416$, $w_4 = -4592$, $w_3 = 3472$, $w_2 = 5096$, $w_1 = -3640$. These weights can be checked by calculating A of Eq. (23) with the model-signal values s_j substituted for the observations v_j. The result should be and is the determinant 23856 of the coefficient matrix. This check is based on the fact that the filter output is proportional to the amplitude of the model-signal content of the input. Another check is based on the fact (discussed in subsection III B) that an LMS filter rejects Maclaurin-series input corresponding to the terms of its clutter model. In this example the filter should and does reject the arbitrary quadratic input

$$[(a + by_j + cy_j^2)] = (1/4)[(4a - 10b + 25c)(4a - 6b + 9c)(4a - 2b + c)$$

$$(4a + 2b + c)(4a + 6b + 9c)(4a + 10b + 25c)].$$

Since any LMS filter blocks constant input, the weights should always sum to zero. This check is part of the previous one, but is noted specially because it is easy to use. The final step is to normalize the weights according to any desired criterion, as discussed in the matched-filter example.

The following points related to solution of Eq. (20) should be noted.

(1) The signal values and observations are equal in number.

(2) The number of observations must equal or exceed the number of model parameters.

(3) If the parameters and observations are equal in number, the best-fit parameters can be calculated from either the equations of condition or the least-squares equations, otherwise the least-squares equations must be used.

(4) With a nonsymmetric signal the clutter model may end with either an even or odd power of y. With a symmetric signal the clutter model must end with an odd power of y.

The second and third points are familiar mathematical facts, the fourth receives additional attention in the next subsection.

(B) Further Analysis and Discussion

This subsection gives further examples, explores the solution of Eq. (20) in greater detail, and discusses the make-up of the filter weights and how the filters operate. Readers with little or no interest in these matters can skim or omit this part. They may want to consider it later, however, if they are puzzled by certain aspects of their calculations for particular filters.

Numerical solutions like the preceding one are easily obtained and are the most efficient way to derive LMS filters for given signals. On the other hand, algebraic solutions provide more information about the properties of the equations and filters. Accordingly, algebraic solutions for centrally symmetric signals will be derived and discussed. These signals receive special attention for two reasons. First, their symmetry introduces certain unique features into the solution of Eq. (20). Second, point-source signals are of this type, and detection of infrared point sources is important in practice.

With the co-ordinate origin at the center of a centrally symmetric signal, all terms containing odd powers of y_j are zero in the coefficient matrix of Eq. (20), and the equations separate into two sets. One set contains the odd clutter parameters, the other the even clutter parameters and the signal amplitude.

$$\begin{bmatrix} s_k^2 & s_k & s_k y_k^2 & s_k y_k^4 \\ s_k & K & y_k^2 & y_k^4 \\ s_k y_k^2 & y_k^2 & y_k^4 & y_k^6 \\ s_k y_k^4 & y_k^4 & y_k^6 & y_k^8 \end{bmatrix} \begin{bmatrix} A \\ B^0 \\ B^2 \\ B^4 \end{bmatrix} = \begin{bmatrix} s_\lambda v_\lambda \\ v_\lambda \\ y_\lambda^2 v_\lambda \\ y_\lambda^4 v_\lambda \end{bmatrix} \tag{26}$$

$$\begin{bmatrix} y_k^2 & y_k^4 & y_k^6 \\ y_k^4 & y_k^6 & y_k^8 \\ y_k^6 & y_k^8 & y_k^{10} \end{bmatrix} \begin{bmatrix} B^1 \\ B^3 \\ B^5 \end{bmatrix} = \begin{bmatrix} y_\lambda v_\lambda \\ y_\lambda^3 v_\lambda \\ y_\lambda^5 v_\lambda \end{bmatrix} \tag{27}$$

The solution of Eq. (26) for the best-fit signal amplitude is

$$A = |s_k^2|(s_\lambda v_\lambda) - |s_k|(v_\lambda) + |s_k y_k^2|(y_\lambda^2 v_\lambda) - |s_k y_k^4|(y_\lambda^4 v_\lambda), \tag{28}$$

where the determinant of the coefficient matrix is ignored because it is absorbed eventually in the filter normalization. The symbols $|s_k^2|$, $|s_k|$, etc. represent the minors of the elements s_k^2, s_k, etc. in the determinant array, e.g.,

$$|s_k^2| = \text{det. of} \begin{bmatrix} K & y_k^2 & y_k^4 \\ y_k^2 & y_k^4 & y_k^6 \\ y_k^4 & y_k^6 & y_k^8 \end{bmatrix}. \tag{29}$$

Since the subscript λ implies summation, Eq. (28) can also be written

$$A = \sum_{j=1}^{K} \{|s_k^2| s_j - |s_k| + |s_k y_k^2| y_j^2 - |s_k y_k^4| y_j^4\} v_j. \tag{30}$$

The subsums in braces are the impulse-response weights of the fifth-order LMS filter for the signal amplitude parameter. Comparison of these expressions with Eq. (B1) of Appendix B shows that the weights are

$$w_{(K+1)-j} = |s_k^2| s_j - |s_k| + |s_k y_k^2| y_j^2 - |s_k y_k^4| y_j^4, \quad j = 1, 2, \ldots K. \tag{31}$$

Least-squares equations applicable with third- and first-degree clutter models can be obtained by eliminating appropriate elements from Eqs. (26,27). Correspondingly, elimination of terms from Eq. (30) gives solutions of the reduced Eq. (26). Specifically, LMS filters based on cubic or linear clutter

models are obtained by eliminating $(|s_k y_k^4|y_j^4)$ or $(|s_k y_k^2|y_j^2 - |s_k y_k^4|y_j^4)$ from Eq. (30). The resulting solutions with the minors expanded are

$$A = \sum_{j=1}^{K} \{[K(y_k^4) - (y_k^2)^2]s_j - [(s_k)(y_k^4) - (y_k^2)(s_k y_k^2)]$$

$$+ [(s_k)(y_k^2) - K(s_k y_k^2)]y_j^2\} \, v_j, \text{ and} \tag{30a}$$

$$A = \sum_{j=1}^{K} \{Ks_j - s_k\} \, v_j. \tag{30b}$$

Again, the subsums in braces are the impulse-response weights and are to be evaluated as previously described. The expanded fifth-order solution is not given because of its length. Specific fifth-order solutions are best obtained directly from Eq. (30).

Table 2 gives algebraic solutions of Eq. (30) for linear, cubic, and quintic models and two symmetric sample patterns. Actually, the sampling is hardly ever symmetric because the signal phase is random. That is usually accounted for, however, by averaging the sample values over the equally probable phases. With a symmetric signal the averages have one of the patterns in Table 2, depending on the ratio of signal length to sampling interval. On average, therefore, the examples of Table 2 include all types of sampled, 1-D, symmetric signals.

No solutions for even-order models are given in Table 2 because none exist. The signal symmetry does not allow the clutter function to end with an even power of y_j. If one attempts to use an even-order model, he finds that the number of observations must be sufficient to permit solution of Eq. (27) for the next higher odd clutter parameter. The fact that this is due to the signal symmetry is not apparent in the least-squares equations, but it can be found in the equations of condition by expanding the coefficient matrix on the row or column containing the signal values.

Table 2 provides examples of several filters for each signal-plus-clutter model. The first example for each model has the least number of observations needed to determine the model parameters. Each successive example then has one more observation than its predecessor. All filter weights are weighted sums of the model signal's even derivatives estimated at the sample point nearest the signal center. This is illustrated by the examples of third-order filters, those based on a third-degree clutter polynomial.

- In the first example, there are just enough samples to estimate the central value of the signal's fourth derivative, and the filter weights are multiples of the estimate $(2b - 8a + 6) \equiv \gamma$.

- In the second example, the number of samples is sufficient to estimate the signal's fifth derivative, but it is zero, and so only the fourth derivative is estimated at the two points nearest the center. The filter weights are multiples of the estimate $(b - 3a + 2) \equiv \delta$.

- In the third example, there are enough samples to estimate central values of both the fourth and sixth derivatives, and the filter weights are weighted sums of the two. However, the derivative values enter the filter weights in different proportions and so do not factor out as they do in the previous examples.

- In the fourth example, the filter weights are weighted sums of the model signal's fourth and sixth derivatives estimated at either point nearest the signal center.

- Presumably, higher and higher even derivatives enter the filter weights as the number of samples is increased.

The same type of comments describe the first- and fifth-order filter weights, except that they are made from different derivatives.

Figure 2 shows that the derivative components of the weights and the terms of the clutter function are interchangeable in a certain sense. For example, with five observations and a linear clutter function, the filter weights contain the second and fourth derivatives. With the same number of observations, adding quadratic and cubic terms to the clutter function eliminates the second derivative from the weights. Further inspection of Fig. 2 provides additional examples of this type. It follows that the weights of Table 2 are built up from values of the model signal's even derivatives above the degree of the clutter function.

Since the filter weights are made from the signal's central derivatives, the weights will be very small or zero if the derivative components are very small or zero. Obviously such filters should be avoided. The cases in Table 2 where the derivative value factors out of the weights may appear to be exceptions. One could argue that only the relative weights matter in such cases, and the derivative factor can be ignored. If the ignored factor is zero, however, the filter output is nevertheless zero with the signal centered on the relative weights. Moreover, since the weights vanish with their derivative components, it can be seen from Eqs. (31,30,28) that in these cases the coefficient matrix of Eq. (26) is singular and there is no filter solution. Thus the filter always vanishes with the model-signal's central derivatives.

From the make-up of the impulse responses, it follows that the filters of Table 2 operate by correlating input with weighted sums of a model-signal's even derivatives; however, the operation can be explained more physically as follows. Let the input at any instant be expanded in a Maclaurin series about the filter center. (This series represents the actual signal and clutter; it is not related to the signal-plus-clutter model.) The filter weights are symmetrically distributed and sum to zero. As a result, the filter eliminates the constant and odd terms of the input series. Furthermore, the weights are so related that even-input terms below the last term of the clutter model also are eliminated. Thus the filter operates by passing only even-input terms higher than the filter order. This explanation can be checked by working out examples from Table 2.

The following useful picture of an LMS filter's operation emerges from the preceding discussion. Let the input consist of additive signal, pure clutter, and random noise, each expanded in a Maclaurin series. The filter blocks input terms that correspond to its clutter model or that have symmetry different from its model signal. The remaining input terms are correlated with combinations of the model-signal's derivatives above the filter order, and the result is passed. The passed quantity is a weighted sum of signal, pure-clutter, and random-noise terms of an order higher than the clutter polynomial. Hence, one can say that an LMS filter assigns to clutter and rejects input (including signal and random noise) that can be represented by the clutter model or has the wrong symmetry, while it assigns to signal and passes input (including pure clutter and random noise) that correlates with the model-signal's derivatives above the filter order.

It is worthwhile to compare 1-D LMS filters for sampled and continuous symmetric signals. The least-squares equations and solutions for the two filter types are very similar, differences being due only to summing or integrating the squared errors over the signal length depending on whether the data are sampled or continuous. Consequently, least-squares equations for continuous data can be obtained from Eqs. (26,27) by substituting signal length $2Y$ for number K of samples per signal length, and replacing sums over samples by integrals over signal length. The same substitutions in Eqs. (30,30a,30b) give the continuous solutions for the best-fit signal-amplitude parameters. For example, with a symmetric signal and a linear clutter model,

$$A = \int_{-Y}^{Y}\{2Ys(y) - \int_{-Y}^{Y} s(\xi)d\xi\}\, v(y)dy = 2Y\int_{-Y}^{Y}\{s(y) - \langle s \rangle\}\, v(y)dy = 2Y\int_{-Y}^{Y} h(-y)v(y)dy. \qquad (32)$$

The first equality is the continuous analog of Eq. (30b). The angular brackets $< >$ in the second equality denote an average over the signal length. Inspection of the last equality gives the continuous impulse response,

$$h(y) = s(-y) - \langle s \rangle, \quad -Y \le y \le Y$$

$$= 0, \quad \text{elsewhere.} \tag{33}$$

From this expression it is easy to show that the continuous impulse response, like the discrete one, is a weighted sum of the model signal's even derivatives above the filter order. Replacing the signal function with its Maclaurin series everywhere in Eq. (33) yields

$$h(y) = \sum_{n=2,4,\ldots}^{\infty} \frac{1}{n!} \left(y^n - \frac{Y^n}{n+1} \right) (D^n s)_0, \tag{34}$$

where $(D^n s)_0$ is the model signal's nth derivative at the signal center. The derivative weights depend on the y co-ordinate in the continuous case and on the sample position in the discrete case. Similar results can be obtained for higher-order filters. All even derivatives above the filter order occur in continuous impulse responses, but discrete impulse responses contain no derivatives beyond the highest that can be estimated with the number of samples available. Decreasing the sampling interval increases the number of samples in the signal length and brings higher derivatives into the filter weights. This indicates that a discrete filter passes smoothly to a continuous one as the sampling interval is reduced to zero.

The behavior described also implies that oversampling should improve LMS-filter performance. Extra samples allow use of a higher-order filter or a more selective lower-order filter whose weights contain a greater number of derivatives. The lower-order filter passes more input terms containing both signal and clutter, and is the better performer if the resulting fractional increase of output signal exceeds that of output clutter. In a staring system for detection of point sources, the needed oversampling requires high spatial resolution, several times higher than is currently available. With a scanning system this limitation can be overcome in the scan direction by increasing the electronic sampling rate. It has been suggested that a similar effect can be achieved in the cross-scan direction by using several columns of detectors with offset centers.[29] It is also possible to increase the number of samples by adding zeros at the ends of the model signal, but that is not as effective as oversampling for two reasons: The zeros provide little extra information about the signal, and the added clutter values may be weakly correlated with those inside the nonzero signal length.

Finally, reference 15 shows that a first-order LMS filter and a matched filter in $(1/f^2)$ clutter have the same transfer function for a raised-cosine signal. This has led to the misconception that an LMS filter is always equivalent to some matched filter. On the contrary, the two types of filters usually are not the same. Relations that determine the circumstances in which they are identical can be obtained by equating general expressions for their impulse responses or transfer functions - e.g., Eqs. (7a) and (33) for 1-D continuous filters and Eqs. (8a) and (31) for 1-D discrete filters. This problem is not studied here because it is only marginally related to computation of filter weights, the main topic of the article.

IV. 2-D LMS FILTERS

The LMS design procedure can be applied in two dimensions almost as easily as in one, although this has not been done heretofore. The derivation of the least-squares equations is the same in both cases; there are simply more equations with more terms in the 2-D case. This makes the 2-D

equations more difficult to solve but does not affect the main features of the solutions. Since the 1-D and 2-D cases are so alike, they are presented in the same way. Subsection IV A contains all the information needed to calculate the weights of a 2-D LMS digital filter for any specified signal. Subsection IV B provides additional examples and discussion of the equations, solutions, and filters for symmetric signals.

(A) Basic Analysis and Example

As in the 1-D case, a model of signal-plus-clutter is fit by least squares to sensor output within the filter, which is the same size as the signal. The signal model is the product of an amplitude A and a shape function s, and the clutter model is an arbitrary continuous function represented within the filter by a few terms of its truncated Maclaurin series. A right-handed co-ordinate system is employed with positive y axis rightward and positive x axis downward (Fig. 3 and Appendix B). To simplify notation, the samples of signal, co-ordinates, and sensor output are labeled with a single ordered subscript. The indexing system is fully specified by the examples in Table 3. With these conventions the signal-plus-clutter model of degree M for K samples can be written

$$f_j = As_j + \sum_{m=0}^{M} \sum_{n=0}^{m} B^{(m-n)n} x_j^{(m-n)} y_j^n$$

$$= As_j + [B^{00} + (B^{10}x_j + B^{01}y_j) + (B^{20}x_j^2 + B^{11}x_jy_j + B^{02}y_j^2)$$

$$+ \ldots (B^{M0}x_j^M + \ldots B^{0M}y_j^M)], \quad j = 1, 2, 3, \ldots K, \text{ and } K \geq \left[1 + \sum_{m=0}^{M} \sum_{n=0}^{m} 1\right]. \tag{35}$$

As before, the superscript of $B^{(m-n)n}$ is not an exponent but simply associates the parameter with the corresponding powers of x_j and y_j. A and $B^{(m-n)n}$ are adjusted to achieve the least-squares fit. In operation the filter algorithm makes these adjustments spontaneously at every point of the scene.

To limit the size of the 2-D least-squares equations, they will be derived for a cubic model ($M = 3$) instead of the quintic model used in the 1-D case. Equation (17) still represents the equations of condition for the differences e_j between model values and observations. Differentiating the sum of the squared differences with respect to the model parameters and equating the derivatives to zero gives

$$e_k \frac{\partial e_k}{\partial A} = 0 \text{ and } e_k \frac{\partial e_k}{\partial [B^{(m-n)n}]} = 0, \quad m = 0, 1, 2, 3, \text{ and } n \leq m. \tag{36}$$

As in Eq. (19), the subscript k implies summation over the K samples in the filter. Equations (17) and (35) are now substituted into Eqs. (36) to obtain the least-squares equations:

$$
\begin{vmatrix}
s_k^2 & s_k & s_k x_k & s_k y_k & s_k x_k^2 & s_k x_k y_k & s_k y_k^2 & s_k x_k^3 & s_k x_k^2 y_k & s_k x_k y_k^2 & s_k y_k^3 \\
s_k & K & x_k & y_k & x_k^2 & x_k y_k & y_k^2 & x_k^3 & x_k^2 y_k & x_k y_k^2 & y_k^3 \\
s_k x_k & x_k & x_k^2 & x_k y_k & x_k^3 & x_k^2 y_k & x_k y_k^2 & x_k^4 & x_k^3 y_k & x_k^2 y_k^2 & x_k y_k^3 \\
s_k y_k & y_k & x_k y_k & y_k^2 & x_k^2 y_k & x_k y_k^2 & y_k^3 & x_k^3 y_k & x_k^2 y_k^2 & x_k y_k^3 & y_k^4 \\
s_k x_k^2 & x_k^2 & x_k^3 & x_k^2 y_k & x_k^4 & x_k^3 y_k & x_k^2 y_k^2 & x_k^5 & x_k^4 y_k & x_k^3 y_k^2 & x_k^2 y_k^3 \\
s_k x_k y_k & x_k y_k & x_k^2 y_k & x_k y_k^2 & x_k^3 y_k & x_k^2 y_k^2 & x_k y_k^3 & x_k^4 y_k & x_k^3 y_k^2 & x_k^2 y_k^3 & x_k y_k^4 \\
s_k y_k^2 & y_k^2 & x_k y_k^2 & y_k^3 & x_k^2 y_k^2 & x_k y_k^3 & y_k^4 & x_k^3 y_k^2 & x_k^2 y_k^3 & x_k y_k^4 & y_k^5 \\
s_k x_k^3 & x_k^3 & x_k^4 & x_k^3 y_k & x_k^5 & x_k^4 y_k & x_k^3 y_k^2 & x_k^6 & x_k^5 y_k & x_k^4 y_k^2 & x_k^3 y_k^3 \\
s_k x_k^2 y_k & x_k^2 y_k & x_k^3 y_k & x_k^2 y_k^2 & x_k^4 y_k & x_k^3 y_k^2 & x_k^2 y_k^3 & x_k^5 y_k & x_k^4 y_k^2 & x_k^3 y_k^3 & x_k^2 y_k^4 \\
s_k x_k y_k^2 & x_k y_k^2 & x_k^2 y_k^2 & x_k y_k^3 & x_k^3 y_k^2 & x_k^2 y_k^3 & x_k y_k^4 & x_k^4 y_k^2 & x_k^3 y_k^3 & x_k^2 y_k^4 & x_k y_k^5 \\
s_k y_k^3 & y_k^3 & x_k y_k^3 & y_k^4 & x_k^2 y_k^3 & x_k y_k^4 & y_k^5 & x_k^3 y_k^3 & x_k^2 y_k^4 & x_k y_k^5 & y_k^6
\end{vmatrix}
\times
\begin{vmatrix}
A \\ B^{00} \\ B^{10} \\ B^{01} \\ B^{20} \\ B^{11} \\ B^{02} \\ B^{30} \\ B^{21} \\ B^{12} \\ B^{03}
\end{vmatrix}
=
\begin{vmatrix}
s_\lambda v_\lambda \\ v_\lambda \\ x_\lambda v_\lambda \\ y_\lambda v_\lambda \\ x_\lambda^2 v_\lambda \\ x_\lambda y_\lambda v_\lambda \\ y_\lambda^2 v_\lambda \\ x_\lambda^3 v_\lambda \\ x_\lambda^2 y_\lambda v_\lambda \\ x_\lambda y_\lambda^2 v_\lambda \\ y_\lambda^3 v_\lambda
\end{vmatrix}
. \tag{37}
$$

In this equation, e.g., $s_k x_k^2 y_k \equiv \sum s_j x_j^2 y_j$, and the dummy index of summation has been changed from k to λ on the right side to facilitate subsequent analysis. Equations for higher-degree models are no more difficult to derive but are inconveniently large. The equations for quartic and quintic models have, respectively, (16 × 16) and (22 × 22) coefficient matrices. Eliminating appropriate elements from Eq. (37) gives equations for lower-degree signal-plus-clutter models. For example, eliminating the last four rows and columns of the coefficient matrix and the last four elements of the parameter and observation vectors gives the least-squares equations for a quadratic model. Unlike 1-D signals, 2-D signals do not always have a central point about which the samples are equally distributed. Even so, the co-ordinate origin should be placed as symmetrically as possible among the samples in order to simplify the coefficient matrix and computation of its elements.

Equation (37) and its reduced forms are applicable to any signal whatsoever. They are solved for the filter weights in the same way as the 1-D Eq. (20). An example will illustrate the procedure. Find the LMS filter for the signal and clutter model below.

Model: $f_j = As_j + [B^{00} + (B^{10}x_j + B^{01}y_j) + (B^{20}x_j^2 + B^{11}x_jy_j + B^{02}y_j^2)]$, $j = 1, 2, 3, \ldots 9$

Signal (s_j)	Indices (j)	Co-ordinates (x_j)	Co-ordinates (y_j)
1 2	1 4	−3/2 −3/2	−1 0
1 0 2	2 5 8	−1/2 −1/2 −1/2	−1 0 1
4 1 3	3 6 9	1/2 1/2 1/2	−1 0 1
2	7	3/2	0

The co-ordinate values result from putting sample 1 in the third quadrant of the x-y plane (Fig. 3), placing the origin between the fifth and sixth samples, and measuring x and y in units of the sampling interval. If the interval is different in the two directions, either interval can be used for the unit of measure, as explained below in subsection IV B. First, Eq. (37) is reduced to its quadratic form, and the elements of the coefficient matrix are calculated from the signal values s_j and co-ordinates (x_j, y_j), with the result

$$\begin{vmatrix} 40 & 16 & 1 & -1 & 14 & 0.5 & 11 \\ 16 & 9 & -1.5 & -1 & 8.25 & 1.5 & 5 \\ 1 & -1.5 & 8.25 & 1.5 & -3.375 & -2.25 & -1.5 \\ -1 & -1 & 1.5 & 5 & -2.25 & -1.5 & -1 \\ 14 & 8.25 & -3.375 & -2.25 & 15.5625 & 3.375 & 3.25 \\ 0.5 & 1.5 & -2.25 & -1.5 & 3.375 & 3.25 & 1.5 \\ 11 & 5 & -1.5 & -1 & 3.25 & 1.5 & 5 \end{vmatrix} \begin{vmatrix} A \\ B^{00} \\ B^{10} \\ B^{01} \\ B^{20} \\ B^{11} \\ B^{02} \end{vmatrix} = \begin{vmatrix} s_\lambda v_\lambda \\ v_\lambda \\ x_\lambda v_\lambda \\ y_\lambda v_\lambda \\ x_\lambda^2 v_\lambda \\ x_\lambda y_\lambda v_\lambda \\ y_\lambda^2 v_\lambda \end{vmatrix} . \quad (38)$$

The solution for the best-fit signal amplitude is

$$A = (7872 s_\lambda v_\lambda - 3300 v_\lambda - 3616 x_\lambda v_\lambda - 336 y_\lambda v_\lambda$$

$$- 5232 x_\lambda^2 v_\lambda + 9888 x_\lambda y_\lambda v_\lambda - 14732 y_\lambda^2 v_\lambda)/28400. \quad (39)$$

The denominator is the determinant of the coefficient matrix, and the numbers in the numerator are, except for signs, the minors of the elements in the first column of the determinant array. As in the 1-D case, the determinant can be omitted since it does not affect the relative filter weights and disappears finally in the normalization. Then, taking account of the summation implied by the subscript and factoring the common multiple v_j from the terms of the numerator gives

$$A = \sum_{j=1}^{9} (7872s_j - 3300 - 3616x_j - 336y_j$$

$$- 5232x_j^2 + 9888x_jy_j - 14736y_j^2)v_j = \sum_{j=1}^{9} w_{(10-j)}v_j. \tag{40}$$

The right side of Eq. (40) can be interpreted as a 2-D convolution sum in which the parenthetical subsums are the LMS filter weights. This is most easily seen if the last equality of Eq. (40) is written in terms of convolution arrays. From Appendix B, Eqs. (B5-B7), and the discussion of those equations,

$$A = \begin{bmatrix} w_9 & w_6 & 0 \\ w_8 & w_5 & w_2 \\ w_7 & w_4 & w_1 \\ 0 & w_3 & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_4 & v_{7a} \\ v_2 & v_5 & v_8 \\ v_3 & v_6 & v_9 \\ v_{3a} & v_7 & v_{9a} \end{bmatrix} = \begin{bmatrix} 0 & w_3 & 0 \\ w_1 & w_4 & w_7 \\ w_2 & w_5 & w_8 \\ 0 & w_6 & w_9 \end{bmatrix} * \begin{bmatrix} v_1 & v_4 & v_{7a} \\ v_2 & v_5 & v_8 \\ v_3 & v_6 & v_9 \\ v_{3a} & v_7 & v_{9a} \end{bmatrix}$$

$$= w_9v_1 + w_8v_2 + w_7v_3 + w_6v_4 + w_5v_5 + w_4v_6 + w_3v_7 + w_2v_8 + w_1v_9. \tag{41}$$

Zero weights and corresponding unused observations (v_{3a}, v_{7a}, v_{9a}) are included in the arrays of Eq. (41) for consistency with the notation of Appendix B. Comparing Eqs. (40) and (41) one finds

$$w_9 = 7872s_1 - 3300 - 3616x_1 - 336y_1$$

$$- 5232x_1^2 + 9888x_1y_1 - 14736y_1^2 = -1344. \tag{42}$$

In the same way the other relative weights are found to be $w_8 = -4384$, $w_7 = 5728$, $w_6 = 6096$, $w_5 = -2800$, $w_4 = 1456$, $w_3 = -4752$, $w_2 = -7072$, $w_1 = 7072$. The weights are checked as in the 1-D case. "A" calculated from Eq. (40) with $v_j = s_j$ should be the determinant 28400 of the coefficient matrix, and the filter should reject arbitrary Maclaurin-series input $[(a + bx_j + cy_j + dx_j^2 + gx_jy_j + ry_j^2)]$ corresponding to the terms of the clutter polynomial. The sum of the weights must be zero for rejection of constant input in the second check. The final step is to normalize the relative weights (set the filter gain) according to any convenient criterion.

Everything needed to calculate an LMS filter for any 2-D signal has now been given. The reader is cautioned, however, that a 2-D LMS filter cannot always be designed simply by solving Eq. (37). With quadratic and higher-degree clutter functions, there is more than one filter for a given signal. This should not be surprising; a matched filter for a given signal also depends on the clutter model, as Eqs. (1a,b) show. The mathematical reasons for the dependence are different in the LMS case, however, and are useful for choosing among alternative LMS filters. This is discussed in the next subsection, where solutions of Eq. (37) for certain symmetric signals are considered.

(B) Further Analysis and Discussion

The signal symmetries of interest are those which cause all terms of the coefficient matrix involving an odd power of x or y to vanish. Examples of these symmetries for linear and higher-degree signal-plus-clutter models are

linear	higher degree
c b d	c b c
a l a	a l a
d b c	c b c

The symmetry for a linear model is slightly lower because Eq. (37) for a linear model contains no coefficients of the type $s_k x_k^p y_k^r$, p and $r \geq 1$. With signals of the specified symmetry, eighty-six matrix elements vanish, and Eq. (37) splits into four equations:

$$\begin{bmatrix} s_k^2 & s_k & s_k x_k^2 & s_k y_k^2 \\ s_k & K & x_k^2 & y_k^2 \\ s_k x_k^2 & x_k^2 & x_k^4 & x_k^2 y_k^2 \\ s_k y_k^2 & y_k^2 & x_k^2 y_k^2 & y_k^4 \end{bmatrix} \begin{bmatrix} A \\ B^{00} \\ B^{20} \\ B^{02} \end{bmatrix} = \begin{bmatrix} s_\lambda v_\lambda \\ v_\lambda \\ x_\lambda^2 v_\lambda \\ y_\lambda^2 v_\lambda \end{bmatrix} \tag{43a}$$

$$(x_k^2 y_k^2) B^{11} = x_\lambda y_\lambda v_\lambda \tag{43b}$$

$$\begin{bmatrix} x_k^2 & x_k^4 & x_k^2 y_k^2 \\ x_k^4 & x_k^6 & x_k^4 y_k^2 \\ x_k^2 y_k^2 & x_k^4 y_k^2 & x_k^2 y_k^4 \end{bmatrix} \begin{bmatrix} B^{10} \\ B^{30} \\ B^{12} \end{bmatrix} = \begin{bmatrix} x_\lambda v_\lambda \\ x_\lambda^3 v_\lambda \\ x_\lambda y_\lambda^2 v_\lambda \end{bmatrix} \tag{43c}$$

$$\begin{bmatrix} y_k^2 & y_k^4 & x_k^2 y_k^2 \\ y_k^4 & y_k^6 & x_k^2 y_k^4 \\ x_k^2 y_k^2 & x_k^2 y_k^4 & x_k^4 y_k^2 \end{bmatrix} \begin{bmatrix} B^{01} \\ B^{03} \\ B^{21} \end{bmatrix} = \begin{bmatrix} y_\lambda v_\lambda \\ y_\lambda^3 v_\lambda \\ x_\lambda^2 y_\lambda v_\lambda \end{bmatrix} \tag{43d}$$

The signal-amplitude solution of Eq. (43a), omitting the determinant, is

$$A = |s_k^2| s_\lambda v_\lambda - |s_k| v_\lambda + |s_k x_k^2| x_\lambda^2 v_\lambda - |s_k y_k^2| y_\lambda^2 v_\lambda \tag{44}$$

$$= \sum_{j=1}^{K} \{ |s_k^2| s_j - |s_k| + |s_k x_k^2| x_j^2 - |s_k y_k^2| y_j^2 \} v_j , \tag{45a}$$

where $|s_k^2|$, $|s_k|$, etc. are the minors of the elements in the first column of the determinant array:

$$|s_k^2| \equiv \text{det. of} \begin{bmatrix} K & x_k^2 & y_k^2 \\ x_k^2 & x_k^4 & x_k^2 y_k^2 \\ y_k^2 & x_k^2 y_k^2 & y_k^4 \end{bmatrix} , \text{ etc.} \tag{46}$$

Equation (45a) gives the best-fit A for cubic signal-plus-clutter models. As in the 1-D case, there are no solutions for quadratic models with signals having the specified symmetry. After elimination of the third and fourth rows and columns from Eq. (43a), the signal-amplitude solution (without determinant) for a linear model is found to be

$$A = \sum_{j=1}^{K} \{Ks_j - s_k\} \, \mathrm{v}_j. \tag{45b}$$

This has the same form as the corresponding 1-D solution Eq. (30b).

The right sides of Eqs. (45a,b) are convolution sums in which the filter weights are the subsums enclosed by braces. Since the impulse response is reversed before convolution with the input v_j, the filter weights for cubic and linear signal-plus-clutter models are respectively

$$w_{(K+1)-j} = |s_k^2|s_j - |s_k| + |s_k x_k^2| x_j^2 - |s_k y_k^2| y_j^2 \text{ and} \tag{47a}$$

$$w_{(K+1)-j} = Ks_j - s_k, \; j = 1, 2, \ldots K. \tag{47b}$$

Expansion of the minors in Eq. (45a) shows that all the terms in braces are equidimensional in x, y, and s. That is, x occurs to the same power in every term, as also do y and s. This means that the relative filter weights are not changed if all values of x_j, or y_j, or s_j are multiplied by a constant, a relation with three notable consequences: (1) The signal and co-ordinates can be scaled to simplify computations if the numbers are awkward. (2) If the x and y sampling intervals are unequal, either can be used as the unit of measure. (3) A difference in the two sampling intervals affects the relative filter weights through the signal values, not through the co-ordinates. Further implications of these facts are pointed out in the last paragraph of this subsection.

Tables 3 and 4 show filters for selected symmetric signals combined with linear and cubic clutter functions. A clutter function is cubic if it has a term $x_j^p y_j^r$ with $(p + r) = 3$. The cubic models were determined by eliminating terms from the complete cubic function until a solution for A containing all values of s_j was found. The complete linear function was used in all cases. Tables 3 and 4 begin with the signals having the fewest samples required to determine the parameters of a linear or cubic model. The other signals were chosen to illustrate the fact that the signal shape, the clutter function, and the number of observations jointly determine the derivative components of the filter weights. In the 1-D examples (Table 2) the number of observations is varied with the clutter function and signal shape fixed. In the 2-D linear examples (Table 3) the clutter function is fixed, but the signal shape changes to maintain the specified symmetry as the number of observations is increased. And in the 2-D cubic examples (Table 4) both the clutter function and the signal shape change with the number of observations. Consequently, the development of the filter weights from the model-signal's derivatives is not illustrated as straightforwardly in two dimensions as in one. The key features are the same in both dimensions, however, and are evident in Tables 2-4. When the number of observations equals the number of model parameters, the filter weights are made from a single model-signal derivative. As the number of observations within the signal is increased, more or different derivatives enter the filter weights depending on the signal shape and clutter function. This suggests, by analogy to the 1-D case, that discrete 2-D filters pass smoothly to continuous form as the sampling interval approaches zero.

Since the filter weights are made from the model-signal's derivatives, the two vanish together. It is also true, as can be seen from Eqs. (47a,45a), that the coefficient matrix in Eq. (43a) is singular when the derivative components of the filter weights are zero. Hence the filter vanishes with the derivatives, even if there is only one component that factors out of the weights, as in the first and fifth examples of Table 4.

The selectivity of the filter depends on the number of derivative components in the weights. The number of observations needed to bring a given number of derivatives into the weights is usually greater in two dimensions than in one because of the greater number of model parameters involved. This makes oversampling of the signal more important for filter performance in two dimensions than in one.

It was noted earlier that an LMS filter, like a matched filter, is not unique but depends on the clutter model as well as the signal. Table 5 exhibits alternative third-order LMS filters for two signals of Table 4. Comparison of the tables shows that clutter terms and model-signal derivatives are interchangeable in the following way. Adding $B^{02}y_j^2$ to the second signal-plus-clutter model of Table 4 eliminates $(D_y^2 s)_4 \sim \beta$ from the filter weights. Similarly, adding $B^{02}y_j^2$ or $B^{20}x_j^2$ to the third model eliminates $(D_y^2 s)_{3,6} \sim \delta$ or $(D_x^2 s)_{3,6} \sim \alpha$ from the filter weights. This effect has at least two consequences. First, in some instances terms must be omitted from the clutter polynomial in order for the filter weights to contain as much information as possible about the signal. That is why the models of Table 4 were determined by eliminating terms from the full cubic clutter function until a signal-amplitude solution containing all signal values was found. Second, the filters of Table 5 reject some input passed by the corresponding filters of Table 4. This is because the filters of Table 5 have more clutter terms and, as pointed out in subsection III B, an LMS filter rejects input representable by its clutter function. Which filters are better depends on whether rejection of additional input, containing both signal and clutter, reduces output signal or output clutter by a greater fraction. The answer probably differs from case to case.

The filter weights and derivative components in Tables 3-5 are for equal x and y sampling intervals. How does changing one of the intervals, say Δx, affect these results? The question is answered by the relations noted earlier between filter weights, co-ordinates, and signal values. With Δy the unit of measure, changing Δx scales x_j values but does not alter the expressions for the relative weights as sums of signal samples a, b, etc.. But the invariance is only algebraic because one or more of the signal values is changed, thereby altering the relative weights numerically. At the same time, derivative components involving Δx are changed by appropriate multiples, and the coefficients of these components in the filter weights are changed inversely by the same multiples, so that the algebraic expressions for the relative weights as sums of signal samples remain unchanged. In practical terms this means that if the definitions of α, β, etc. are substituted into the relative filter weights in Tables 3-5, the resulting expressions for the weights in terms of a, b, etc. are valid for any Δx and Δy. (In fact the weights were derived as sums of signal values and then, to show their basic structure, were converted to the tabulated sums of derivative values with $\Delta x = \Delta y = 1$ sample interval.)

V. SUMMARY

Spatial filters are often used to enhance signal-to-clutter ratio in an electro-optic sensor's output prior to signal detection. Matched filters and LMS filters are two types commonly employed. A matched filter maximizes the ratio of filtered-signal energy to average filtered-clutter power. An LMS filter implements a regression analysis of the sensor output and estimates the amplitude of the signal present.

This article gives methods of computing the impulse-response weights of one- and two-dimensional digital matched filters and LMS filters for any signal and clutter representable by certain types of models. The methods and models are applicable to outputs from scanning or staring optical sensors and to signals from extended or point sources. The signal at the sensor output must be known to design either kind of filter. With resolved signals, whose shapes are variable, the design techniques are limited to cases where an average signal is adequate or a particular form of the signal must be detected. For matched-filter design the clutter power spectrum at the sensor output also must be known, but for LMS design prior knowledge of the background is not essential, though it may be

helpful for choosing the clutter model. Images of sensor output are the best sources of the required information about signals and backgrounds.

The matched-filter design technique is developed by using isotropic power-spectral clutter models, Eqs. (4a,b), which contain only even powers of spatial frequency. This leads to fairly simple expressions, Eqs. (7a-c), for the 1-D and 2-D continuous impulse responses as weighted sums of the model signals and their derivatives. The weights are the coefficients of terms in the power-spectral models. Consequently, these coefficients must be known either theoretically or from background measurements, preferably the latter. The 1-D impulse-response equation contains only even derivatives; the 2-D equation contains only powers of the Laplacian operator $[(\partial^2/\partial x^2) + (\partial^2/\partial y^2)]^p$, $p = 1, 2, 3$. These equations are converted to discrete form by numerical approximation of the derivatives. The results are a pair of operator equations (8a,b) for the filter weights and a table of derivative operator arrays (Table 1). 2-D arrays are tabulated only for equal x and y sampling intervals. The derivations of the equations and operators are given in sufficient detail so that a reader can repeat them with different clutter models and unequal x and y sampling intervals. If the clutter model contains odd powers of frequency, the impulse response is not easily determined. For that reason, such models were not used in this work.

LMS filters are derived from least-squares fits of deterministic signal-plus-clutter models, Eqs. (16, 35) to sensor output. The clutter models employed here are truncated and sampled Maclaurin-series expansions of continuous functions. Their coefficients, unlike those of the power-spectral models, need not be known because an LMS filter constantly and automatically adjusts the parameters of its signal-plus-clutter model to suit the input, as in Fig. 1. However, information about the background is useful for choosing the clutter function since it is not specified by LMS design, which rather makes best use of the model chosen. Least-squares equations (20) are given for 1-D LMS filters based on first- through fifth-degree clutter polynomials. Even the fifth-degree equations are easily solved with a hand-held calculator capable of inverting a (7×7) matrix. The equations for 2-D LMS filters are much like those for 1-D filters. There are simply more of them, and each has more terms than its 1-D counterpart. Least-squares equations (37) are given for 2-D LMS filters based on first- through third-degree clutter models. Solution of the third-degree equations requires inversion of an (11×11) matrix. Fourth- and fifth-degree equations are not difficult to derive or to solve numerically with a computer, but they are not given because they involve (16×16) and (22×22) coefficient matrices.

LMS filters for specific signals are most easily obtained by solving the least-squares equations numerically, but the solutions provide little understanding of the filters. For understanding, algebraic solutions are needed. Accordingly, the equations are solved algebraically for symmetric signals, which are of special interest and present special problems. Only two types of symmetric, 1-D, sampled signals are possible on average, and solutions are given for both (Table 2). In the 2-D case, however, the treatment is necessarily limited to a few of the many possible symmetric signals (Tables 3-5).

Examination of the algebraic solutions shows that LMS digital-filter weights are sums of the model-signal's derivatives estimated as near as possible to the signal center. Likewise, continuous LMS impulse responses can be expressed as sums of the model-signal's centrally evaluated derivatives, e.g., Eq. (34). The derivative components of the impulse response and the terms of the clutter model are interchangeable; that is, adding a term to the clutter model eliminates a derivative from the impulse response (Fig. 2 and Table 5). A continuous LMS impulse response contains all signal derivatives not ruled out by the clutter model. The derivative components of a discrete LMS impulse response are further limited by the number of signal samples. A derivative cannot appear in the filter weights if there are not enough samples to estimate it. As the sampling interval is decreased, the number of samples in the signal span increases, and more of the signal's derivatives enter the filter weights (Fig. 2). This shows how a discrete LMS filter passes to continuous form with decrease of

sampling interval. Improvement of filter performance by oversampling also is indicated since a filter's selectivity depends on the number of derivatives in its weights.

The build-up of the impulse response from the model-signal's derivatives, except those ruled out by the clutter model, suggests two simple explanations of an LMS filter's operation. Clearly the filter correlates input with the signal derivatives, but a more physical explanation can be given as follows. An LMS filter assigns input to signal, clutter, and residual. Assigned clutter is all input representable by the clutter model. Assigned signal is the remaining input that correlates with the model signal. Residual is any input left over (none if the model parameters and observations are equal in number). Assigned signal, clutter, and residual all contain actual signal and clutter in different proportions. Assigned clutter and residual are rejected, assigned signal is passed. Addition of a term to the clutter model causes additional portions of actual signal and clutter to be assigned to clutter and/or residual and rejected. Whether performance improves depends on whether this causes actual signal or actual clutter in the output to decrease by the greater fraction, a question that must be answered case-by-case. Notice that this physical explanation essentially describes what a regression analysis does, in keeping with the LMS filter's mathematical basis.

Table 1 — Derivative Operators and Array Approximations[a]

Operator	Array Symbol and Approximation
(d/dy)	$[w^1] = [1\ -1]$ or $\frac{1}{2}[1\ 0\ -1]$[b]
(d^2/dy^2)	$[w^2] = [1\ -2\ 1]$
(d^3/dy^3)	$[w^3] = [1\ -3\ 3\ -1]$ or $\frac{1}{2}[1\ -2\ 0\ 2\ -1]$[b]
(d^4/dy^4)	$[w^4] = [1\ -4\ 6\ -4\ 1]$
(d^5/dy^5)	$[w^5] = [1\ -5\ 10\ -10\ 5\ -1]$ or $\frac{1}{2}[1\ -4\ 5\ 0\ -5\ 4\ -1]$[b]
(d^6/dy^6)	$[w^6] = [1\ -6\ 15\ -20\ 15\ -6\ 1]$
$(\partial^2/\partial y^2)$	$[w^{02}] = [w^2] = [1\ -2\ 1]$
$(\partial^2/\partial x^2)$	$[w^{20}] = [w^2]^t$
$\nabla^2 = (\partial^2/\partial x^2) + (\partial^2/\partial y^2)$	$[\nabla^2] = [w^{20}] + [w^{02}] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
$(\partial^4/\partial y^4)$	$[w^{04}] = [w^4] = [1\ -4\ 6\ -4\ 1]$
$(\partial^4/\partial x^4)$	$[w^{40}] = [w^4]^t$
$(\partial^4/\partial x^2 \partial y^2)$	$[w^{22}] = [w^2]^t [w^2] = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$
$\nabla^4 = \nabla^2 \nabla^2$	$[\nabla^4] = [w^{40}] + 2[w^{22}] + [w^{04}]$
	$= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$
$(\partial^6/\partial y^6)$	$[w^{06}] = [w^6] = [1\ -6\ 15\ -20\ 15\ -6\ 1]$
$(\partial^6/\partial x^6)$	$[w^{60}] = [w^6]^t$
$(\partial^6/\partial x^2 \partial y^4)$	$[w^{24}] = [w^2]^t [w^4] = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ -2 & 8 & -12 & 8 & -2 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}$
$(\partial^6/\partial x^4 \partial y^2)$	$[w^{42}] = [w^4]^t [w^2] = [w^{24}]^t$
$\nabla^6 = \nabla^2 \nabla^4$	$[\nabla^6] = [w^{60}] + 3[w^{42}] + 3[w^{24}] + [w^{06}]$
	$= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3 & -12 & 3 & 0 & 0 \\ 0 & 3 & -24 & 57 & -24 & 3 & 0 \\ 1 & -12 & 57 & -112 & 57 & -12 & 1 \\ 0 & 3 & -24 & 57 & -24 & 3 & 0 \\ 0 & 0 & 3 & -12 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

a. The meaning and use of these arrays are discussed in Appendices B and C. They are not vectors or matrices except where the symbol t for "transpose" occurs. In that case they are regarded as row vectors or as matrices for evaluating the expression involved. The arrays operate by convolution, indicated by an asterisk *. The first step of convolution is array reversal. Convolving an operator array with an array of samples estimates a derivative at the sample points. Operator arrays must always fully overlap sample arrays. This table gives operator arrays for $\Delta y = \Delta x = 1$ sample interval. 2-D arrays for unequal x and y sampling intervals must be constructed as described in Appendix C.

b. The first array is from forward and backward difference approximations of the derivative. The second array from a central difference approximation. The two approximations are the same for even order derivatives. Central arrays estimate derivatives at central weights; forward and backward arrays at first and last weights. Even-order arrays involve only even operators; they are effectively derived from central difference formulae and thus estimate derivatives at central weights.

Table 2 — 1-D LMS Filters for Symmetric Signals

Signal + Clutter Model	No of Obs (K)	Model Signal Values (s_i)*	Indices (i) of Sample Points	Signal Derivative Components of Filter Weights†	Relative Weights of Signal Amplitude Filter
$As_i + \sum_{m=0}^{J} B^m y_i^m$ 3 parameters	3	a 1 a	1 2 3	$(D^2s)_2 \sim (2a\ 2) \equiv \alpha$	$\alpha\{1\ \ 2\ 1\}^{\S}$
	4	a 1 1 a	1 2 3 4	$(D^2s)_{2,3} \sim (a\ 1) \equiv \beta$	$\beta\{1\ \ 1\ \ 1\ 1\}^{\S}$
	5	b a 1 a b	1 2 3 4 5	$(D^2s)_3 \sim \alpha$ $(D^4s)_3 \sim (2b\ \ 8a+6) \equiv \gamma$	$(10\alpha+3\gamma)(\ 5\alpha\ \ 2\gamma)(-10\alpha\ \ 2\gamma)(-5\alpha\ \ 2\gamma)(10\alpha+3\gamma)$
	6	b a 1 1 a b	1 2 3 4 5 6	$(D^2s)_{3,4} \sim \beta$ $(D^4s)_{3,4} \sim (b\ \ 3a+2) \equiv \delta$	$(5\beta+2\delta)(-\beta\ \ \delta)(\ 4\beta\ \ \delta)(-4\beta-\delta)(-\beta\ \ \delta)(5\beta+2\delta)$
	7	c b a 1 a b c	1 2 3 4 5 6 7	$(D^2s)_4 \sim \alpha$ $(D^4s)_4 \sim \gamma$ $(D^6s)_4 \sim (2c\ \ 12b+30a\ \ 20) \equiv \iota$	$(35\alpha+28\gamma+5\iota)(-7\gamma\ \ 2\iota)(-21\alpha\ \ 14\gamma-2\iota)(-28\alpha\ \ 14\gamma\ \ 2\iota)$ etc.[‖]
$As_i + \sum_{m=0}^{J} B^m y_i^m$ 5 parameters	5	b a 1 a b	1 2 3 4 5	$(D^4s)_3 \sim \gamma$	$\gamma\{1\ \ 4\ 6\ \ 4\ 1\}^{\S}$
	6	b a 1 1 a b	1 2 3 4 5 6	$(D^4s)_{3,4} \sim \delta$	$\delta\{1\ \ 3\ 2\ 2\ \ 3\ 1\}^{\S}$
	7	c b a 1 a b c	1 2 3 4 5 6 7	$(D^4s)_4 \sim \gamma$ $(D^6s)_4 \sim \iota$	$(18\gamma+5\iota)(-42\gamma\ \ 12\iota)(6\gamma+3\iota)(36\gamma+8\iota)(6\gamma+3\iota)(-42\gamma\ \ 12\iota)(18\gamma+5\iota)$
	8	c b a 1 1 a b c	1 2 3 4 5 6 7 8	$(D^4s)_{4,5} \sim \delta$ $(D^6s)_{4,5} \sim (c\ \ 5b+9a\ \ 5) \equiv \zeta$	$(21\delta+7\zeta)(-39\delta-14\zeta)(-9\delta)(27\delta+7\zeta)(27\delta+7\zeta)(-9\delta)(-39\delta-14\zeta)(21\delta+7\zeta)$
$As_i + \sum_{m=0}^{J} B^m y_i^m$ 7 parameters	7	c b a 1 a b c	1 2 3 4 5 6 7	$(D^6s)_4 \sim \iota$	$\iota\{1\ \ -6\ 15\ \ -20\ 15\ \ -6\ 1\}^{\S}$
	8	c b a 1 1 a b c	1 2 3 4 5 6 7 8	$(D^6s)_{4,5} \sim \zeta$	$\zeta\{1\ \ -5\ 9\ \ -5\ \ -5\ 9\ \ -5\ 1\}^{\S}$
	9	d c b a 1 a b c d	1 2 3 4 5 6 7 8 9	$(D^6s)_6 \sim \iota$ $(D^8s)_5 \sim (2d-16c+56b-112a+70) \equiv \theta$	$(52\iota+14\theta)(-221\iota-60\theta)(286\iota+80\theta)(13\iota-4\theta)(-260\iota-80\theta)$ etc.[‖]
	10	d c b a 1 1 a b c d	1 2 3 4 5 6 7 8 9 10	$(D^6s)_{5,6} \sim \zeta$ $(D^8s)_{5,6} \sim (d\ \ 7c+20b\ \ 28a+14) \equiv \phi$	$(429\zeta+132\phi)(-1573\zeta-495\phi)(1430\zeta+495\phi)(858\zeta+165\phi)(-1144\zeta-297\phi)$ etc.[‖]

*a b c d may have any values except those which make the derivative components vanish. The central symmetry of the signals ensures that $\Sigma\, s_i y_i^m = \Sigma\, y_i^m = 0$ for odd m. These are the conditions for Eq. (20) to reduce to Eqs. (26,27).

†$(D^m s)_{i,j} \sim (d^m s\ \ dy^m)$ at sample point i or j. The derivatives are estimated from Eqs. (C2–C7) of Appendix C. If the derivative components vanish, the coefficient matrix of Eq. (26) is singular, and the equation cannot be solved for a signal amplitude filter.

‡The relative filter weights are in braces. Ordinarily the derivative factor is an unseen part of the filter normalization (gain) constant. The derivative factor is shown here to emphasize that the weights are built up from the model signal's derivatives estimated as near as possible to the signal center, and that the filter vanishes with the derivative components.

‖The remaining weights are centrally symmetric with the preceding ones.

Table 3 — 2-D, First-order, LMS Filters for Certain Symmetric Signals

Signal + Clutter Model	No. of Obs. (K)	Model Signal (s_j)*	Indices of Sample Points (j)	Signal-Derivative Components of Filter Weights+	Relative Weights of Signal-Amplitude Filter&
–	4	a 1 1 a	1 3 2 4	$(D_xD_ys)_{1,4}\sim(2a-2)\equiv\alpha$	$\alpha\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}$ §
	5	b a 1 a b	2 1 3 5 4	$(D_y^2s)_3\sim\alpha$ $(D_x^2s)_3\sim(2b-2)\equiv\beta$	$(-2\alpha+3\beta)$ $(3\alpha-2\beta)\ (-2\alpha-2\beta)\ (3\alpha-2\beta)$ $(-2\alpha+3\beta)$
	6	a b 1 1 b a	1 4 2 5 3 6	$(D_xD_ys)_{2,5}\sim(a-b)\equiv\gamma$ $(D_x^2s)_{2,5}\sim(a-2+b)\equiv\delta$	$(3\gamma+\delta)\ (-3\gamma+\delta)$ $(-2\delta)\ (-2\delta)$ $(-3\gamma+\delta)\ (3\gamma+\delta)$
$As_j+(B^{00}+B^{10}x_j+B^{01}y_j)$ – 4 parameters –	7	c b a 1 a b c	2 3 1 4 7 5 6	$(D_y^2s)_4\sim\alpha$ $(D_x^2s)_4\sim\beta$ $(D_x^4s)_4\sim(2c-6b+6)\equiv\varepsilon$	$(-2\alpha+18\beta+5\varepsilon)$ $(-2\alpha-3\beta-2\varepsilon)$ $(5\alpha-10\beta-2\varepsilon)\ (-2\alpha-10\beta-2\varepsilon)\ (5\alpha-10\beta-2\varepsilon)$ $(-2\alpha-3\beta-2\varepsilon)$ $(-2\alpha+18\beta+5\varepsilon)$
	8	b c a 1 1 a c b	1 5 2 6 3 7 4 8	$(D_xD_ys)_{2,7}\sim\alpha$ $(D_x^2s)_{2,7}\sim(b-2a+1)\equiv\zeta$ $(D_x^2s)_{3,6}\sim(c-2+a)\equiv\theta$ $(D_x^3D_ys)_{1,8}\sim(-2c+2b-6a+6)\equiv\sigma$	$(3\alpha+\zeta+\theta+\sigma)\ (-3\alpha+\zeta+\theta-\sigma)$ $(\alpha-\zeta-\theta)\ (-\alpha-\zeta-\theta)$ $(-\alpha-\zeta-\theta)\ (\alpha-\zeta-\theta)$ $(-3\alpha+\zeta+\theta-\sigma)\ (3\alpha+\zeta+\theta+\sigma)$
	9	c b d a 1 a d b c	1 4 7 2 5 8 3 6 9	$(D_y^2s)_5\sim\alpha$ $(D_x^2s)_5\sim\beta$ $(D_xD_ys)_{1,5,9}\sim(1-a-b+c)\equiv\phi$ $(D_xD_ys)_{2,4,6,8}\sim(-1+a+b-d)\equiv\psi$	$(3\alpha+3\beta+14\phi+4\psi)\ (-6\alpha+3\beta-4\phi+4\psi)\ (3\alpha+3\beta-4\phi-14\psi)$ $(3\alpha-6\beta-4\phi+4\psi)\ (-6\alpha-6\beta-4\phi+4\psi)\ (3\alpha-6\beta-4\phi-4\psi)$ $(3\alpha+3\beta-4\phi-14\psi)\ (-6\alpha+3\beta-4\phi+4\psi)\ (3\alpha+3\beta+14\phi+4\psi)$

* a,b,c,d may have any values except those which make the derivative components vanish. The symmetry of the signals is such that $\Sigma\, sx = \Sigma\, sy = \Sigma\, x = \Sigma\, y = \Sigma\, x_jy_j = 0$. These are the conditions for Eq. (45b) to be a solution of Eq. (37).

\+ $(D_x^pD_y^qs)_{i,j} = \partial^{(p+q)}s/\partial x^p\partial y^q$ at sample point i or j. The derivatives are estimated from Eqs. (C2-C7) of Appendix C with $\Delta x = \Delta y = 1$ sample interval. If the derivative components vanish, the coefficient matrix of Eq. (43a) is singular, and the equation cannot be solved for a signal-amplitude filter.

& In these examples the relative weights are given also by $(Ks_j - \Sigma s_j)$. The conditions for this are (1) a linear clutter function $(B^{00}+B^{10}x_j+B^{01}y_j)$ and (2) the signal symmetry specified above.

§ The relative filter weights are in braces. Ordinarily the derivative factor is an unseen part of the filter normalization (gain) constant. The derivative factor is shown here to emphasize that the weights are built up from the model-signal's derivatives estimated as near as possible to the signal center and that the filter vanishes with the derivative components.

Table 4 — 2-D, Third-order, LMS Filters for Certain Symmetric Signals

Signal + Clutter Model	No. of Obs. (K)	Model Signal (a_j)*	Indices of Sample Points (j)	Signal-Derivative Components of Filter Weights†	Relative Weights of Signal-Amplitude Filter
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{11}x_jy_j+B^{21}x_j^2y_j)$ — 6 parameters —	6	a a 1 1 a a	1 4 2 5 3 6	$(D_x^2s)_{2,5}\sim(2a-2)\equiv\alpha$	$\alpha\begin{Bmatrix}1 & 1\\ -2 & -2\\ 1 & 1\end{Bmatrix}$‡
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{30}x_j^3)$ — 6 parameters —	7	b a c 1 c a b	2 3 1 4 7 5 6	$(D_y^2s)_4\sim(2c-2)\equiv\beta$ $(D_x^4s)_4\sim(2b-8a+6)\equiv\gamma$	$(6\beta+3\gamma)$ $(-24\beta-12\gamma)$ $(35\beta+6\gamma)\ (-34\beta+6\gamma)\ (35\beta+6\gamma)$ $(-24\beta-12\gamma)$ $(6\beta+3\gamma)$
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{11}x_jy_j+B^{21}x_j^2y_j+B^{03}y_j^3)$ — 7 parameters —	8	a a b 1 1 b a a	2 5 1 3 6 8 4 7	$(D_x^2s)_{3,6}\sim\alpha$ $(D_y^2s)_{3,6}\sim(b-1)\equiv\delta$	$(2\alpha-2\delta)\ (2\alpha-2\delta)$ $(-2\alpha+6\delta)\ (-2\alpha-2\delta)\ (-2\alpha-2\delta)\ (-2\alpha+6\delta)$ $(2\alpha-2\delta)\ (2\alpha-2\delta)$
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{02}y_j^2+B^{30}x_j^3+B^{03}y_j^3)$ — 8 parameters —	9	b a d c 1 c d a b	3 4 1 2 5 8 9 6 7	$(D_x^4s)_5\sim\gamma$ $(D_y^4s)_5\sim(2d-8c+6)\equiv\epsilon$	$(35\gamma-18\epsilon)$ $(-140\gamma+72\epsilon)$ $(-18\gamma+35\epsilon)\ (72\gamma-140\epsilon)\ (102\gamma+102\epsilon)\ (72\gamma-140\epsilon)\ (-18\gamma+35\epsilon)$ $(-140\gamma+72\epsilon)$ $(35\gamma-18\epsilon)$
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{11}x_jy_j+B^{02}y_j^2$ $+B^{21}x_j^2y_j+B^{12}x_jy_j^2)$ — 9 parameters —	9	b a b c 1 c b a b	1 4 7 2 5 8 3 6 9	$(D_x^2D_y^2s)_5\sim(-4c+4b-4a+4)\equiv\zeta$	$\zeta\begin{Bmatrix}1 & -2 & 1\\ -2 & 4 & -2\\ 1 & -2 & 1\end{Bmatrix}$‡
$Aa_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{11}x_jy_j+B^{02}y_j^2$ $+B^{21}x_j^2y_j+B^{12}x_jy_j^2+B^{03}y_j^3)$ — 10 parameters —	11	b a b d c 1 c d b a b	2 5 8 1 3 6 9 11 4 7 10	$(D_y^4s)_6\sim\epsilon$ $(D_x^2D_y^2s)_6\sim\zeta$	$(-20\epsilon+35\zeta)\ (40\epsilon-70\zeta)\ (-20\epsilon+35\zeta)$ $(18\epsilon-20\zeta)\ (-32\epsilon+10\zeta)\ (28\epsilon+20\zeta)\ (-32\epsilon+10\zeta)\ (18\epsilon-20\zeta)$ $(-20\epsilon+35\zeta)\ (40\epsilon-70\zeta)\ (-20\epsilon+35\zeta)$

* a,b,c,d may have any values except those which make the derivative components vanish. The symmetry of the signals is such that $\Sigma\, s_i x_i^p y_i^r = \Sigma\, x_i^p y_i^r = 0$ with p, or r, or both odd. These are the conditions for Eq. (37) to reduce to Eqs. (43a,b,c,d).

† $(D_x^pD_y^rs)_{i,j} = \partial^{p+r}s/\partial x^p\partial y^r$ at sample point i or j. The derivatives are estimated from Eqs.(C2-C7) of Appendix C with $\Delta x = \Delta y = 1$ sample interval. If the derivative components vanish, the coefficient matrix of Eq. (43a) is singular, and the equation cannot be solved for a signal-amplitude filter.

‡ The relative filter weights are in braces. Ordinarily the derivative factor is an unseen part of the filter normalization (gain) constant. The derivative factor is shown here to emphasize that the weights are built up from the model-signal's derivatives estimated as near as possible to the signal center, and that the filter vanishes with the derivative components.

Table 5 — Alternative Third-order LMS Filters for Two Signals of Table 4

Signal + Clutter Model	No. of Obs. (K)	Model Signal (s_j)*	Indices of Sample Points (j)	Signal-Derivative Components of Filter Weights+	Relative Weights of Signal-Amplitude Filter
$As_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{02}y_j^2+B^{30}x_j^3)$ −7 parameters−	7	b a c 1 c a b	2 3 1 4 7 5 6	$(D_x^4 s)_4 \sim (2b-8a+6) \equiv \gamma$	$\gamma \begin{Bmatrix} & 1 & \\ & -4 & \\ 0 & 6 & 0 \\ & -4 & \\ & 1 & \end{Bmatrix}$ $
$As_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{11}x_jy_j+B^{02}y_j^2+B^{21}x_j^2y_j+B^{03}y_j^3)$ −8 parameters−	8	a a b 1 1 b a a	2 5 1 3 6 8 4 7	$(D_x^2 s)_{3,6} \sim (2a-2) \equiv \alpha$	$\alpha \begin{Bmatrix} 1 & 1 \\ 0 -2 -2 & 0 \\ 1 & 1 \end{Bmatrix}$ $
$As_j+(B^{00}+B^{10}x_j+B^{01}y_j$ $+B^{20}x_j^2+B^{11}x_jy_j+B^{21}x_j^2y_j+B^{03}y_j^3)$				$(D_y^2 s)_{3,6} \sim (b-1) \equiv \delta$	$\delta \begin{Bmatrix} 0 & 0 \\ 1 -1 -1 & 1 \\ 0 & 0 \end{Bmatrix}$ $

*a,b,c may have any values except those which make the derivative components vanish. The signal symmetry is such that $\Sigma\, s_j x_j^p y_j^r = \Sigma\, x_j^p y_j^r = 0$ with p, or r, or both odd. These are the conditions for Eq. (37) to reduce to Eqs. (43 a,b,c,d).

+$(D_x^p D_y^r s)_{i,j} = \partial^{(p+r)} s / \partial x^p \partial y^r$ at sample point i or j. The derivatives are estimated from Eqs. (C2-C7) of Appendix C with $\Delta x = \Delta y = 1$ sample interval. If the derivative components vanish, the coefficient matrix of Eq. (43a) is singular, and the equation cannot be solved for a signal-amplitude filter.

$The relative filter weights are in braces. Ordinarily the derivative factor is an unseen part of the filter normalization (gain) constant. The derivative factor is shown here to emphasize that the weights are built up from the model-signal's derivatives evaluated as near as possible to the signal center, and that the filter vanishes with the derivative components. In practice the zero weights are not included in the impulse response because they do not affect the filter output. They are included here to show that the calculations actually give zero weights at the indicated points.

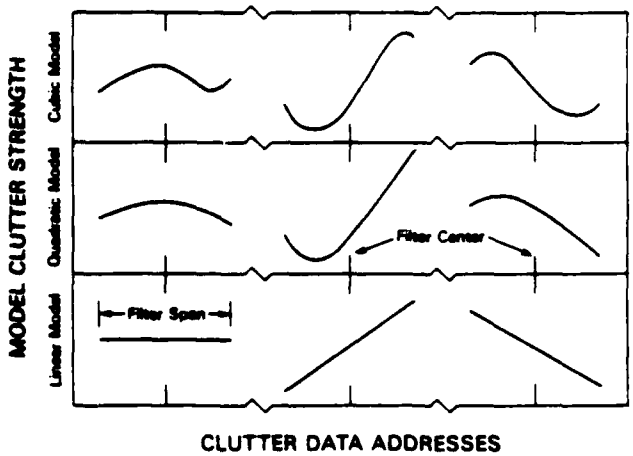


Fig. 1 — Adjustment of 1-D clutter models to local background. This figure is discussed in the first paragraph of subsection III A.

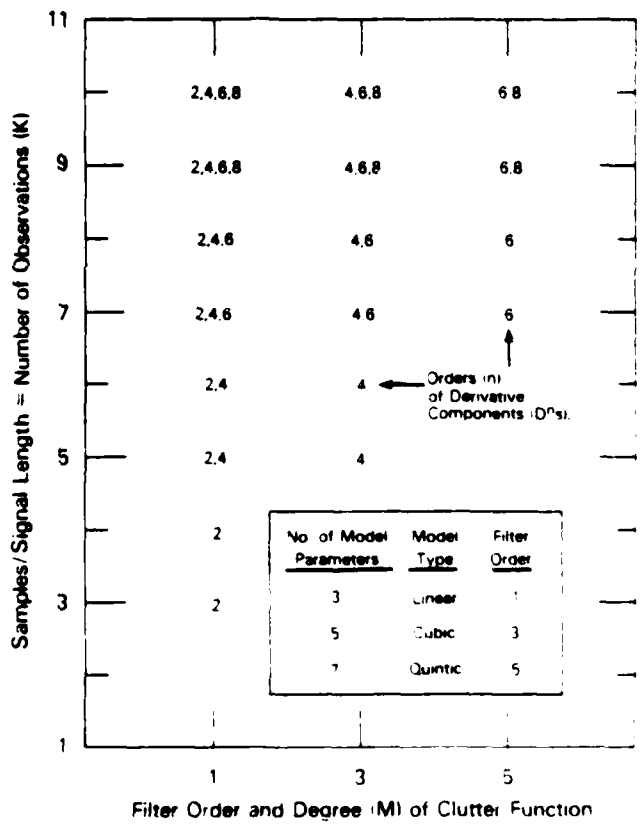


Fig. 2 — Signal-derivative components $(D^n)_{i,j}$ of 1-D, LMS, digital filters for symmetric signals. This figure shows the model-signal derivatives that make up the weights of first-, third-, and fifth-order filters with increasing number of samples/signal length, i.e., decreasing sampling interval. Comparison of columns shows that the weights are made of derivatives above the degree of the clutter model (filter order). Reading up the columns, one sees that more and higher derivatives enter the weights as the sampling interval is decreased. Reading across the rows, one sees that adding terms to the clutter polynomial eliminates model-signal derivatives from the filter weights.

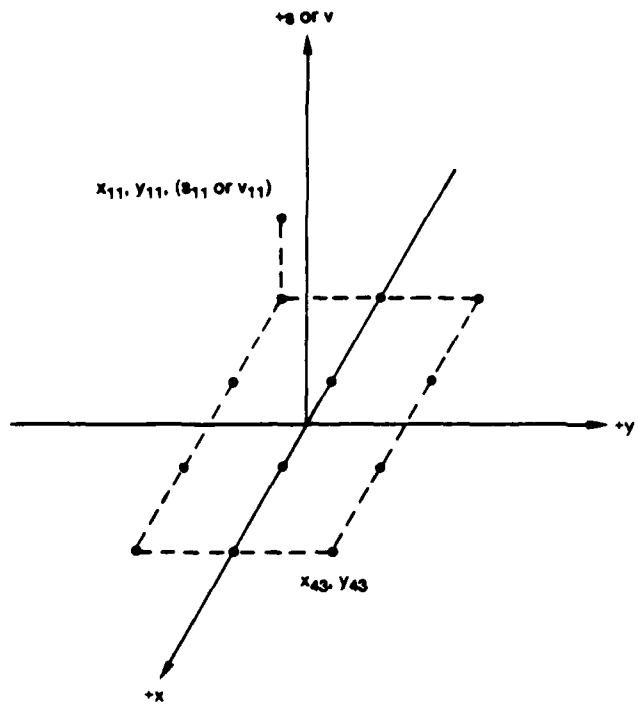


Fig. 3 — Co-ordinate system and numbering of array rows and columns. "s" is model-signal strength; "v" is observed sensor output from actual signal and clutter. This figure is discussed in the last paragraph of Appendix B.

APPENDIX A

Clutter Models

The clutter models used in the derivations of LMS filters are deterministic, i.e., represent unaveraged local clutter properties. These models are simply continuous functions that are fit by least squares to the background within the filters. In such limited regions the models can be represented by a few terms of their Maclaurin series Eqs. (16 and 35). Maclaurin series are used for simplicity; expansions in any complete set of orthogonal functions also could be used. Whatever the expansion, the deterministic models conform to the input; i.e., the least-squares values of the expansion coefficients change as the background passes through the LMS filter.

Statistical models are used for the matched-filter derivations. These models are fixed; they do not respond to local background variations; rather, their parameters are adjusted before use to approximate certain globally averaged properties of the anticipated clutter. The remainder of this appendix discusses three candidate statistical models and the reasons for choosing one of them.

Temporal stationarity and spatial isotropy are important qualities related to the choice. A process is stationary and isotropic if its average properties depend only on time and distance between its points (ref. 27, sec. 10.3.2). Most types of optical clutter are naturally nonstationary and anisotropic. Moreover, these qualities are affected by experimental conditions. Nonstationarity is increased if the sensor views different backgrounds in rapid succession. Anisotropy may be induced or altered by the sensor's construction, and is always affected by the sensor's perspective because the length scale varies with direction in the background plane unless the scene is viewed along a normal.[11] Nevertheless, clutter often is represented by stationary isotropic models for reasons of convenience and ignorance: Nonstationarity and anisotropy are mathematically difficult to treat, and too little is known about them to make the effort worthwhile. In addition, unless a certain anisotropy predominates because of the sensor's construction, clutter will tend to be isotropic on average if the sensor views many scenes from all perspectives.[12,13]

Clutter usually is modeled in the time and space domain by autocorrelation functions, and in the frequency domain by their Fourier transforms, power spectra. In this article time dependence and hence stationarity are ignored. Three purely spatial autocorrelation functions that have been proposed as models are[9-13]

$$R(\Delta r) = \sigma^2 \exp(-|\Delta r|/L), \tag{A1}$$

$$R(\Delta x, \gamma\Delta y) = \sigma^2 \exp[-(\Delta x^2 + \gamma^2\Delta y^2)^{1/2}/L], \text{ and} \tag{A2}$$

$$R(\Delta x, \Delta y) = \sigma^2 \exp[-(|\Delta x|/L_x) - (|\Delta y|/L_y)]. \tag{A3}$$

Here $\Delta r = \Delta x + \Delta y$ is a displacement vector, σ^2 is clutter variance, L is correlation length, and γ is a factor that accounts for a change of scale due to perspective. The first model is isotropic; the second is anisotropic due to perspective alone; the third is intrinsically anisotropic, i.e., without

regard to perspective and even if $L_x = L_y$. The power spectra derived from these autocorrelation functions are[11]

$$N(\Omega^2) = F[R(\Delta r)] = 2\pi\sigma^2 L^2/(1 + L^2\Omega^2)^{3/2}, \tag{A4}$$

$$N(u^2, v^2/\gamma^2) = F[R(\Delta x, \gamma\Delta y)] = 2\pi\sigma^2(L^2/\gamma)/[1 + L^2u^2 + L^2(v/\gamma)^2]^{3/2}, \text{ and} \tag{A5}$$

$$N(u^2, v^2) = F[R(\Delta x, \Delta y)] = 4\sigma^2 L_x L_y/[(1 + L_x^2u^2)(1 + L_y^2v^2)], \tag{A6}$$

where F indicates Fourier transformation and $\hat{\Omega} = \hat{u} + \hat{v}$ is 2-D spatial frequency with x and y components $\hat{u}$, $\hat{v}$. Note that Eq. (A4) is a function of the combination $\Omega^2 = (u^2 + v^2)$, not a function of the separate variables u^2, v^2. This results from the isotropy and consequent circular symmetry of the autocorrelation function, Eq. (A1) (ref. 11, app. B; ref. 28, pp. 244-248). A 1-D power spectral model is obtained heuristically from Eq. (A4) by substituting $u = 0$ and replacing v by ω for notational clarity.

For reasons already explained, the isotropic power spectral model is used here. However, Eq. (A4) and its 1-D counterpart are mathematically awkward for deriving matched filters. More suitable forms are obtained by expanding the denominators in Maclaurin series. This leads to

$$N(\omega^2) = N_0/(1 + aL^2\omega^2 + bL^4\omega^4 + cL^6\omega^6) \text{ and} \tag{A7}$$

$$N(\Omega^2) = N_0/(1 + aL^2\Omega^2 + bL^4\Omega^4 + cL^6\Omega^6), \tag{A8}$$

where N_0 is power-spectral density at zero frequency. Terms beyond the fourth are dropped because they lead to filters larger than any expected correlation lengths. (N_0, aL^2, bL^4, cL^6) can be regarded as disposable parameters to be determined from measurements of natural backgrounds. In reference 30 an approximation to a 1-D power spectrum is derived, rather than heuristically inferred, from Eq. (A4); the approximate function varies as $1/(1 + d\omega^2)$. Additional terms in the denominator of Eq. (A7) can be regarded as empirical higher-order approximations. Equation (A8) is an isotropic model because it depends only on $\Omega = +\sqrt{u^2 + v^2}$ and therefore its transform, the autocorrelation function, depends only on the magnitude of $\hat{\Delta r}$ (ref. 28, pp. 244-248).

An isotropic power spectrum like Eq. (A8), but containing both odd and even powers of Ω, can be obtained by expanding the denominator of $N(\Omega) = N_0/f(\Omega)$ in a Maclaurin series:

$$N(\Omega) = N_0/(1 + \delta\Omega + \epsilon\Omega^2 + \zeta\Omega^3 + \ldots\ldots) \tag{A9}$$

This model can decrease less rapidly than $(1/\Omega^2)$ as some measured power spectra do. Since physical quantities are real, however, autocorrelation functions and hence power spectra must be real and even. This ordinarily is considered to rule out odd powers of frequency in series or polynomial power-spectral models (ref. 31, sec. 5.10). But if such models are functions of frequency magnitude, they can contain odd powers and yet be even. Thus models with odd powers of frequency seem physically permissible. When Eq. (A9) is substituted into Eq. (2b), however, the odd powers lead to terms in the impulse response which cannot be related simply and generally to the signal. This disadvantage of the odd powers is considered to outweigh the greater flexibility they provide for fitting measurements. For that reason models with odd powers of frequency are not used in this article.

Attention has been directed to certain consistencies and discrepancies between physical properties of clutter and mathematical properties of statistical models. Though such comparisons are significant and useful, it should be remembered that the Wiener-Khintchine type of models employed here

are primarily mathematical constructs and tools. They were not discovered by analysis of physical processes, and their connection with natural phenomena often is obscure.[32] Thus their suitability is to be judged more by results obtained with them than by physical consonance or dissonance between model and phenomenon.

APPENDIX B

Convolution Arrays and Co-ordinates

The operation of nonrecursive digital filters can be represented by tables of sampled input data, impulse-response weights, and output values. In Appendix C it is shown that the same type of tables can be used to represent numerical approximations of derivative operators. For reasons that will become apparent, these tables are called convolution arrays. Reference 28 (pp. 30-40) explains the basis of this representation and an alternative using vectors and matrices. The vector and matrix representation is discussed at length in references 8,33,34. Convolution arrays are not vectors or matrices because they do not multiply as vectors and matrices do; the multiplication rule is described below. However, addition and multiplication by a scalar are performed as with vectors and matrices. To distinguish the two types of quantities, convolution arrays will be enclosed in brackets [ ], vectors and matrices in parentheses. Convolution arrays lack the mathematical completeness and power of vectors and matrices but are much simpler. In this article, convolution arrays are used wherever possible.

The elements of a 1-D convolution array are written in a row and labeled with indices j. For example,

$$[\mathrm{v}_1\ \mathrm{v}_2\ \mathrm{v}_3 \ldots \mathrm{v}_j\ \mathrm{v}_{j+1}\ \mathrm{v}_{j+2} \ldots]\ \text{and}\ [w_1\ w_2\ w_3],$$

with v_j and w_j being input values and impulse-response weights respectively. An input and impulse response are convolved by rotating the weight array 180° about its center, passing it stepwise over the input array, and at each step summing the products of coincident elements. Graphically, a single step of the process is

$$\begin{array}{l} [w_3\ w_2\ w_1] \rightarrow \\ \quad x \quad x \quad x \\ [\mathrm{v}_1\ \mathrm{v}_2\ \mathrm{v}_3 \ldots \mathrm{v}_j\ \mathrm{v}_{j+1}\ \mathrm{v}_{j+2} \ldots] = \sum_{\beta=0}^{2} w_{3-\beta}\mathrm{v}_{j+\beta}. \end{array} \tag{B1}$$

Rotating the weight array 180° reverses one of the sequences as convolution requires. Either array could be rotated, but if the input sequence is reversed, some thought may be necessary to obtain the output in the right order. Reversing the weight array avoids the need for thought. A quantity like that on the right side of Eq. (B1), in which one index decreases while the other increases, is called a convolution sum since it approximates a single value of a convolution integral. Passing the weight array over the input gives a point-by-point numeric estimate of the convolution integral's functional form. The resulting filter output is a row of convolution sums (ref. 28, p. 31). With $*$ denoting convolution, the output obtained by passing a five-element input through a two-weight filter is

$$[w_1\ w_2]*[\mathrm{v}_1\ \mathrm{v}_2\ \mathrm{v}_3\ \mathrm{v}_4\ \mathrm{v}_5] = \tag{B2}$$

$$[(w_2\mathrm{v}_1 + w_1\mathrm{v}_2)(w_2\mathrm{v}_2 + w_1\mathrm{v}_3)(w_2\mathrm{v}_3 + w_1\mathrm{v}_4)(w_2\mathrm{v}_4 + w_1\mathrm{v}_5)].$$

The number n of input samples must equal or exceed the number m of weights, and the weights must always fully overlap the input. Consequently, the output has $(m - 1)$ fewer values than the input, i.e., $n - (m - 1)$ values. Sometimes the input can be extended by adding peripheral zeros, but only when it is known to be zero beyond the listed samples. If the input above were known to begin and end with zeros, then the filter output would be

$$[w_1\ w_2] * [0\ v_1\ v_2\ v_3\ v_4\ v_5\ 0] = \tag{B3}$$

$$[(w_1v_1)(w_2v_1 + w_1v_2)(w_2v_2 + w_1v_3)(w_2v_3 + w_1v_4)(w_2v_4 + w_1v_5)(w_2v_5)].$$

If two arrays are the same size, they can be multiplied by placing one over the other without rotating either, and then summing products of coincident elements. Multiplication is indicated simply by writing the arrays in sequence. Thus

$$[v_1\ v_2\ v_3][w_1\ w_2\ w_3] = w_1v_1 + w_2v_2 + w_3v_3. \tag{B4}$$

The elements of a two-dimensional convolution array are written in square or rectangular form and can be indexed in various ways. A common method is to label them with row and column numbers ij, for example,

$$\begin{bmatrix} v_{11} & v_{12} & \cdot & \cdot \\ v_{21} & v_{22} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & v_{ij} \end{bmatrix} \text{ and } \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}.$$

This input and impulse response are convolved by rotating the weight array 180° about its center, passing it stepwise horizontally and vertically over the input array, and at each step summing products of coincident elements. Graphically, a single step of the process is

$$\begin{bmatrix} w_{22} & w_{21} \\ w_{12} & w_{11} \end{bmatrix} \begin{matrix} x & x \\ x & x \end{matrix} \quad \begin{bmatrix} v_{11} & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & v_{ij} & v_{i(j+1)} & \cdots \\ \cdots & v_{(i+1)j} & v_{(i+1)(j+1)} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & v_{kn} \end{bmatrix} = \sum_{\alpha=0}^{1} \sum_{\beta=0}^{1} w_{(2-\alpha)(2-\beta)} v_{(i+\alpha)(j+\beta)}. \tag{B5}$$

Here also, rotating the weight array 180° about its center reverses it as convolution requires. The quantity on the right side of Eq. (B5) is a 2-D convolution sum, and the filter output is an array of such sums. The output from a four-element filter with a nine-element input is

$$\begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} * \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} = \tag{B6}$$

$$\begin{bmatrix} (w_{22}v_{11} + w_{21}v_{12} + w_{12}v_{21} + w_{11}v_{22}) & (w_{22}v_{12} + w_{21}v_{13} + w_{12}v_{22} + w_{11}v_{23}) \\ (w_{22}v_{21} + w_{21}v_{22} + w_{12}v_{31} + w_{11}v_{32}) & (w_{22}v_{22} + w_{21}v_{23} + w_{12}v_{32} + w_{11}v_{33}) \end{bmatrix}.$$

As in the 1-D case, the two arrays being convolved must always fully overlap. Thus, if the impulse response has m rows and n columns, the output array has $(m - 1)$ fewer rows and $(n - 1)$ fewer columns than the input array. Peripheral zeros may be added to extend 2-D input arrays only if they are known to be zero beyond the listed samples. The procedure for multiplying 2-D arrays of equal size is analogous to the procedure for 1-D arrays.

Elements of 2-D arrays also can be labeled with a single ordered subscript. Such a system for matrices is described in references 8,33,34. In that system, labeling begins at the top left element of the array, goes down the column, then to the top of the next column, and continues thus to the bottom right element. Labeling of course does not affect the rules for addition, multiplication by a scalar, multiplication, or convolution. For example,

$$\begin{bmatrix} w_1 & w_4 & w_7 \\ w_2 & w_5 & w_8 \\ w_3 & w_6 & w_9 \end{bmatrix} \begin{bmatrix} v_1 & v_4 & v_7 \\ v_2 & v_5 & v_8 \\ v_3 & v_6 & v_9 \end{bmatrix} = \sum_{i=1}^{9} w_i v_i. \tag{B7}$$

This is not the only possible system of labeling with a single ordered subscript. The purpose of such a system is to simplify a discussion or derivation by simplifying notation, or to extend 1-D mathematical techniques to two dimensions. Any system that does either is permissible.

Input data in a convolution array are related as well to a co-ordinate system. Consequently, the connection between the arrangement of the data in the array and in the co-ordinate system should be specified. The 3-D co-ordinate system used here is shown in Fig. 3. It is a right-handed system with the positive x and y directions downward and rightward respectively. This choice of co-ordinates reconciles two conventional notations: x corresponds to row number, y to column number; x precedes y, and row number precedes column number; x and row number increase downward, y and column number increase rightward. To simplify computations, the co-ordinate origin is placed as symmetrically as possible among the samples of signal. With the usual rectangular sampling grid, the origin is located at the center of the smallest rectangle enclosing the signal, and the unit of measure is the sampling interval in either the x or y direction. If the two intervals are the same, as in Fig. 3, the co-ordinates of the sample points are integers or odd half integers. Row and column numbering begins with 11 at the third-quadrant corner of the smallest rectangle enclosing the signal. Alternatively, the co-ordinates, signals, and observations can be labeled with a single subscript in any convenient order, as already explained.

APPENDIX C

Approximation of Derivatives and Derivative Operators

This appendix describes estimation of derivatives by finite differences and approximation of derivative operators by convolution arrays. The finite-difference formulae are needed to identify derivative components of LMS-filter weights and to derive the operator arrays, which are needed to calculate matched-filter weights from Eqs. (8a,b).

1. Finite-Difference Approximations

The basic relation between finite differences and derivatives is (ref. 35, art. 18)

$$\frac{\Delta^n f(y)}{\Delta y^n} = f^{(n)}(y + \theta n \Delta y), \quad 0 < \theta < 1. \tag{C1}$$

This relation says that if an interval is divided into n equal subintervals of length Δy, the nth difference of a continuous function $f(y)$ in the interval, divided by Δy^n, is equal to the nth derivative of the function somewhere in the interval. Equation (C1) is not completely satisfactory for applications because it does not specify the point where the derivative is evaluated. However, most numerical-analysis texts derive more specific relations from interpolation formulae, e.g., the Newton-Sterling central-difference formula, and the Gregory-Newton forward- and backward-difference formulae.[35–43] These formulae yield the following lowest-order approximations of the first six derivatives (ref. 35, art. 48; ref. 36, secs. 35,36).

$$\begin{aligned}
\Delta y(dv/dy) &\cong \Delta v_1 = \nabla v_2 = -v_1 + v_2 \\
\Delta y^2(d^2v/dy^2) &\cong \Delta^2 v_1 = \nabla^2 v_3 = \delta^2 v_2 = v_1 - 2v_2 + v_3 \\
\Delta y^3(d^3v/dy^3) &\cong \Delta^3 v_1 = \nabla^3 v_4 = -v_1 + 3v_2 - 3v_3 + v_4 \\
\Delta y^4(d^4v/dy^4) &\cong \Delta^4 v_1 = \nabla^4 v_5 = \delta^4 v_3 = v_1 - 4v_2 + 6v_3 - 4v_4 + v_5 \\
\Delta y^5(d^5v/dy^5) &\cong \Delta^5 v_1 = \nabla^5 v_6 = -v_1 + 5v_2 - 10v_3 + 10v_4 - 5v_5 + v_6 \\
\Delta y^6(d^6v/dy^6) &\cong \Delta^6 v_1 = \nabla^6 v_7 = \delta^6 v_4 = v_1 - 6v_2 + 15v_3 - 20v_4 + 15v_5 - 6v_6 + v_7
\end{aligned} \tag{C2-C7}$$

Here y is position or time, v is sensor output, v_j is a sample of sensor output, and Δ^n, ∇^n, δ^n are forward, backward, and central differences. The forward-, backward-, and central-difference formulae are the same for even derivatives, but for odd derivatives the lowest-order central-difference approximations are

$$\begin{aligned}
\Delta y(dv/dy) &\cong \mu\delta v_2 = \frac{1}{2}(-v_1 + v_3), \\
\Delta y^3(d^3v/dy^3) &\cong \mu\delta^3 v_3 = \frac{1}{2}(-v_1 + 2v_2 - 2v_4 + v_5), \\
\Delta y^5(d^5v/dy^5) &\cong \mu\delta^5 v_4 = \frac{1}{2}(-v_1 + 4v_2 - 5v_3 + 5v_5 - 4v_6 + v_7).
\end{aligned} \tag{C8-C10}$$

If the unit of y is one sample interval, $\Delta y^n = 1$ and the formulae as written approximate the derivatives themselves. The central-difference formulae are usually more accurate than the forward- and backward-difference formulae (ref. 36, sec. 29). The formulae for the fifth and sixth derivatives are included for completeness; ordinarily they are not used because differences higher than the fourth tend to be erratic (ref. 35, art. 17).

As noted, these formulae are the lowest-order approximations of the derivatives. Higher-order approximations are available but are seldom used for filters, possibly because additional samples are required, and reduced correlation between the added and original samples may offset the greater accuracy of the higher-order formulae. In addition, the higher-order approximations require higher-order differences, which usually are avoided as already explained. Second-order central-difference approximations of the first six derivatives are

$$\Delta y(dv/dy) \cong \mu\delta v_3 - \frac{1}{12}\mu\delta^3 v_3 = \frac{1}{12}(v_1 - 8v_2 + 8v_4 - v_5),$$

$$\Delta y^2(d^2v/dy^2) \cong \delta^2 v_3 - \frac{1}{12}\delta^4 v_3 = \frac{1}{12}(-v_1 + 16v_2 - 30v_3 + 16v_4 - v_5),$$

$$\Delta y^3(d^3v/dy^3) \cong \mu\delta^3 v_4 - \frac{1}{4}\mu\delta^5 v_4 = \frac{1}{8}(v_1 - 8v_2 + 13v_3 - 13v_5 + 8v_6 - v_7),$$

$$\Delta y^4(d^4v/dy^4) \cong \delta^4 v_4 - \frac{1}{6}\delta^6 v_4 = \frac{1}{6}(-v_1 + 12v_2 - 39v_3 + 56v_4 - 39v_5 + 12v_6 - v_7), \qquad \text{(C11-C16)}$$

$$\Delta y^5(d^5v/dy^5) \cong \mu\delta^5 v_5 - \frac{1}{3}\mu\delta^7 v_5 = \frac{1}{6}(v_1 - 9v_2 + 26v_3 - 29v_4 + 29v_6 - 26v_7 + 9v_8 - v_9),$$

$$\Delta y^6(d^6v/dy^6) \cong \delta^6 v_5 - \frac{1}{4}\delta^8 v_5 = \frac{1}{4}(-v_1 + 12v_2 - 52v_3 + 116v_4 - 150v_5 + 116v_6 - 52v_7 + 12v_8 - v_9).$$

Extensive tables containing coefficients of v_j in formulae like Eqs. (C2-C16) can be found in reference 44. These tables provide many additional approximations with estimates of the errors involved in their use.

Equations (C2-C16) approximate derivatives at different sample points from sets of values beginning with the same sample v_1. The forward-, backward-, and central-difference formulae estimate the derivatives, respectively, at the first (v_1), last (v_2 to v_7), and central (v_2, v_3, v_4, v_5) sample points. This must be taken into account when approximating partial derivatives and sums of derivatives. There are numerous alternatives to these formulae as indicated by the tables of reference 44. That should not be surprising since this is a problem in approximation, and such problems usually can be solved in many ways. In particular, the derivatives need not be estimated only at first, last, and central points or even at sample points; the interpolation formulae permit their estimation at arbitrary points. However, Eqs. (C2-C7) are most satisfactory for the purposes of this article. If other formulae are used, it is sometimes impossible to identify the derivative components of the LMS filter weights. The probable cause of this is pointed out in the next part.

2. *Derivative-Operator Arrays*

Arrays approximating continuous derivative operators will now be obtained by writing Eqs. (C2-C10) in terms of convolution arrays. It is important to remember that the finite-difference equations determine how and where the derivatives are estimated. The arrays do not affect the approximations; they are simply a convenient way of writing the equations.

With $\Delta y = 1$ sample interval, the lowest-order central-difference formula for the first derivative is

$$dv/dy \cong \frac{1}{2}[-1\ 0\ 1][v_1\ v_2\ v_3] = \frac{1}{2}(-v_1 + v_3).$$

The row of weights approximates the operator (d/dy), but it is not very useful because it can operate only once on an array of its own size. Its utility is much increased by reversing the weights to form a convolution array and writing

$$(d/dy) \cong \frac{1}{2}[1\ 0\ -1].$$

This array, when convolved with any other of equal or greater size, operates repeatedly on the other array to estimate the derivative at the sample points that coincide with the array's central weight, i.e.,

$$\frac{1}{2}[1\ 0\ -1] * [\ldots\ v_{n-1} v_n v_{n+1} v_{n+2} \ldots] \cong [\ldots\ (dv/dy)_n\ (dv/dy)_{n+1} \ldots].$$

Thus both the derivative's values and its functional form are approximated numerically. Convolution arrays approximating the other monovariate-derivative operators can be constructed in the same way. The results are given in Table 1 where the array for the nth-derivative operator is represented by $[w^n]$. Since these operators are actually zero beyond the listed weights, they can be extended by peripheral zeros when necessary.

Before taking up partial derivatives, it is necessary to consider some properties of the 1-D operator arrays and implications of those properties. It may seem that the result of convolving discrete operators should be predictable by analogy to continuous operators. That is true for discrete operators derived from Eqs. (C2-C7)—the forward-, backward-, and even central-difference formulae. For example, convolving $[w^1]$ and $[w^2]$ from Eqs. (C2) and (C3) gives $[w^3]$ from Eq. (C4), as expected by analogy to $(d/dy)(d^2/dy^2) = (d^3/dy^3)$:

$$[w^1] * [w^2] = [1\ -1] * [0\ 1\ -2\ 1\ 0] = [1\ -3\ 3\ -1] = [w^3].$$

But operators derived from Eqs. (C8-C16)—the odd central-difference and higher-order formulae—do not behave likewise. This emphasizes that discrete operators represent finite-difference formulae and must be used consistently with them, not by analogy to continuous operators. Conversely, the discrete-operator behavior indicates that Eqs. (C2-C7) are analytically related in the same way as continuous derivatives, but Eqs. (C8-C16) are not. This explains why the former equations are more satisfactory than the latter for identifying the derivative components of LMS-filter weights. In applications where numeric accuracy is more important than analytic properties, Eqs. (C8-C16) should be better.

Turning now to bivariate partial derivatives, the associated operators are approximated by square or rectangular arrays whose horizontal and vertical directions correspond, respectively, to the x and y variables (Fig. 3). The first step in constructing a bivariate operator is to derive an approximation for the partial derivative from finite-difference equations. The monovariate parent equations determine the point where the partial derivative is estimated. Monovariate operators are then combined to form

a 2-D array that acts on data to reproduce the approximate formula.[45] The procedure is best explained with examples. Initially it will be assumed that $\Delta x = \Delta y = 1$ sample interval. Frequently, however, the sampling interval is different in the two directions, and the difference must be taken into account as described at the end of this appendix.

Central-difference operators for even homogeneous derivatives are easily constructed. For example, from Eq. (C3)

$$(\partial^2 v/\partial y^2) \cong v_{x1} - 2v_{x2} + v_{x3} \text{ and } (\partial^2 v/\partial x^2) \cong v_{1y} - 2v_{2y} + v_{3y}, \tag{C17}$$

with the derivatives estimated at the central points. Now suppose we have the data array

$$\begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix},$$

and we want to estimate $(\partial^2 v/\partial y^2)$ and $(\partial^2 v/\partial x^2)$ at the position of v_{22}. Add peripheral zeros to $[w^2]$ of Table 1 and to its transpose $[w^2]^t$, and write tentatively

$$(\partial^2/\partial y^2) \cong \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} = [w^{02}], \quad (\partial^2/\partial x^2) \cong \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = [w^{20}].$$

Next, apply these operators to the data array, and compare the results with Eqs. (C17).

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} = v_{21} - 2v_{22} + v_{23} \cong (\partial^2 v/\partial y^2)_{22}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} = v_{12} - 2v_{22} + v_{32} \cong (\partial^2 v/\partial x^2)_{22}$$

The operators give the desired derivative estimates. Since the estimates are made at the same point and the arrays are equidimensional, the operators can be added to obtain the approximate Laplacian operator.

$$\nabla^2 = (\partial^2/\partial y^2) + (\partial^2/\partial x^2) \cong [w^{02}] + [w^{20}] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} = [\nabla^2].$$

It should be noted that the peripheral zeros are not essential in the arrays for $(\partial^2/\partial y^2)$ and $(\partial^2/\partial x^2)$, which also can be written

$$(\partial^2/\partial y^2) \cong [1 \ -2 \ 1] \text{ and } (\partial^2/\partial x^2) \cong \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

However, these operators cannot be added to obtain the Laplacian operator since they are not equidimensional and do not necessarily estimate the derivatives at the same point. Convolving any of these

operators with a data array of equal or greater size estimates a derivative or sum of derivatives at sample points that coincide with the operator center. For example,

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} v_{11} & \cdots & v_{1j} \\ \vdots & \cdots & \vdots \\ v_{i1} & \cdots & v_{ij} \end{bmatrix} \cong \begin{bmatrix} ? & \cdots & ? \\ \vdots & (\nabla^2 v)_{22} \cdots (\nabla^2 v)_{2(j-1)} & \vdots \\ \vdots & (\nabla^2 v)_{(i-1)2} \cdots (\nabla^2 v)_{(i-1)(j-1)} & \vdots \\ ? & \cdots & ? \end{bmatrix}.$$

The question marks indicate that the Laplacian derivative is indeterminate at the sample points in the first and last rows and columns because the operator, when centered at these points, lies partly outside the data array.

Forward- and backward-difference operators are constructed in the same way as central-difference operators. For example, from Eq. (C2)

$$(\partial v/\partial y) \cong -v_{x1} + v_{x2} \text{ and } (\partial v/\partial x) \cong -v_{1y} + v_{2y}, \tag{C18}$$

with the derivatives estimated at either the first or last sample points. From $[w^1]$ and $[w^1]^t$ of Table 1 we have tentatively

$$(\partial/\partial y) \cong [1\ -1] = [w^{01}] \text{ and } (\partial/\partial x) \cong \begin{bmatrix} 1 \\ -1 \end{bmatrix} = [w^{10}].$$

To determine if and where these operators approximate the derivatives, apply them to a (2×2) data array, and compare the results with Eqs. (C18).

$$[1\ -1] * \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} (-v_{11} + v_{12}) \\ (-v_{21} + v_{22}) \end{bmatrix} \cong \begin{bmatrix} (\partial v/\partial y)_{11} \\ (\partial v/\partial y)_{21} \end{bmatrix} \cong \begin{bmatrix} (\partial v/\partial y)_{12} \\ (\partial v/\partial y)_{22} \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} * \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = [(-v_{11} + v_{21})(-v_{12} + v_{22})] \cong [(\partial v/\partial x)_{11}\ (\partial v/\partial x)_{12})]$$

$$\cong [(\partial v/\partial x)_{21}\ (\partial v/\partial x)_{22}]$$

The operators do indeed estimate the derivatives at the indicated points. The positions of the estimates are ambiguous, but the ambiguity is in the forward- and backward-difference formulae. The arrays correctly represent those formulae.

Operators for mixed partial derivatives are constructed by the same methods used for homogeneous derivatives. Let us find the forward- and backward-difference operator for $(\partial^3 v/\partial x \partial y^2)$. From Eqs. (C2,C3)

$$(\partial^2 v/\partial y^2) \cong v_{x1} - 2v_{x2} + v_{x3},$$

$$(\partial^3 v/\partial x \partial y^2) = (\partial/\partial x)(\partial^2 v/\partial y^2) \cong (-v_{11} + v_{21}) - 2(-v_{12} + v_{22}) + (-v_{13} + v_{23}). \tag{C19}$$

with the third derivative estimated at either v_{11} or v_{23}. The data array that contains only the samples in Eq. (C19) is

$$\begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \end{bmatrix}.$$

An array of numbers is needed which, when reversed and multiplied with this data array, gives the right side of Eq. (C19). By inspection the required array is

$$[w^{12}] = \begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \end{bmatrix}.$$

If the 1-D forward- and backward-difference operators $[w^1]$, $[w^2]$ are regarded temporarily as row vectors, the array above is the matrix product $[w^1]^t[w^2]$. Thus the desired 2-D forward- and backward-difference operator is

$$(\partial^3/\partial x\,\partial y^2) \cong [w^1]^t[w^2] = [w^{12}].$$

The derivative is estimated at either the top-left or lower-right corner of the array, the ambiguity again originating in the forward- and backward-difference formulae.

Next let us find the central-difference operator for the same derivative. From Eqs. (C3,C8)

$$(\partial^2 v/\partial y^2) \cong v_{x1} - 2v_{x2} + v_{x3},$$

$$(\partial/\partial x)(\partial^2 v/\partial y^2) \cong \frac{1}{2}(-v_{11} + v_{31}) - (-v_{12} + v_{32}) + \frac{1}{2}(-v_{13} + v_{33}), \qquad \text{(C20)}$$

with the third derivative estimated at the position of v_{22}. The right side of Eq. (C20) in array notation is

$$\frac{1}{2}\begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix} * \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix},$$

so the first array must be an operator which approximates $(\partial^3 v/\partial x\,\partial y^2)$ at the array center. This array is the matrix product of the 1-D central-difference operators $[w^1]^t$ and $[w^2]$ regarded temporarily as row vectors. Thus the desired 2-D central-difference operator is

$$(\partial^3/\partial x\,\partial y^2) \cong [w^1]^t[w^2] = [w^{12}].$$

This is the same as the relation for the corresponding 2-D forward- and backward-difference operator.

In general, for any mixed partial derivative,

$$(\partial^{m+n}/\partial x^m\,\partial y^n) \cong [w^m]^t[w^n] = [w^{mn}],$$

where $[w^m]$, $[w^n]$ are 1-D operators regarded temporarily as row vectors, and t denotes the transpose. The 1-D operators may represent either central-difference formulae or forward- and backward-difference formulae. The formulae determine the points where $[w^{mn}]$ estimates the derivative.

Thus far it has been assumed that $\Delta y = \Delta x = 1$ sample interval. If the sample interval is not the same in the two directions, the finite-difference formulae must be expressed in a common unit before the operator arrays are constructed. Suppose the Laplacian-operator array with $\Delta x = 3\Delta y$ is required. Either sample interval can be chosen as the unit, but to avoid fractions in the arrays, let Δx be the unit. Then from Eq. (C3)

$$\frac{1}{9}(\partial^2 v/\partial y^2) \cong v_{x1} - 2v_{x2} + v_3 \text{ and } (\partial^2 v/\partial x^2) \cong v_{1y} - 2v_{2y} + v_{3y}.$$

The corresponding operator arrays are

$$(\partial^2/\partial y^2) \cong \begin{bmatrix} 0 & 0 & 0 \\ 9 & -18 & 9 \\ 0 & 0 & 0 \end{bmatrix} = [w^{02}] \text{ and } (\partial^2/\partial x^2) \cong \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = [w^{20}].$$

Since these arrays are equidimensional, have the same units, and estimate the derivatives at the same point, they can be added to obtain the Laplacian-operator array for sampled data with $\Delta x = 3\Delta y$.

$$[\nabla^2] = [w^{02}] + [w^{20}] = \begin{bmatrix} 0 & 1 & 0 \\ 9 & -20 & 9 \\ 0 & 1 & 0 \end{bmatrix}$$

It should be noted that use of a common unit for the x and y variables does not resolve all issues raised by unequal x and y sampling intervals. In those circumstances, for example, the Nyquist frequency and the accuracy of the derivative estimates are different in the two directions. What do these facts imply? Such issues are worth considering, but they do not affect calculation of filter weights and, therefore, are not addressed here.

REFERENCES

1. D.O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems," Tech. Rept. No. PTR-6C (RCA Laboratories, Princeton, N.J., 1943). Reprinted in Proc. IEEE **51**, 1016 (1963).

2. G.L. Turin, "An introduction to matched filters," IRE Trans. Inf. Theory **IT-6**, 311 (1960).

3. E.C. Westerfield, R.H. Prager, and J.L. Stewart, "Processing gains against reverberation (clutter) using matched filters," IRE Trans. Inf. Theory **IT-6**, 342 (1960).

4. D. Middleton, "On a new class of matched filters and generalizations of the matched filter concept," IRE Trans. Inf. Theory **IT-6**, 349 (1960).

5. L.J. Cutrona, E.N. Leith, C.J. Palermo, and L.J. Porcello, "Optical data processing and filtering systems," IRE Trans. Inf. Theory **IT-6**, 386 (1960).

6. G.L. Turin, "An introduction to digital matched filters," Proc. IEEE **64**, 1092 (1976).

7. H.C. Andrews, *Computer Techniques in Image Processing* (Academic Press, New York, 1970).

8. W.K. Pratt, *Digital Image Processing* (Wiley, New York, 1978).

9. J.J. Otazo and R.R. Parenti, "Digital filters for the detection of resolved and unresolved targets embedded in infrared (IR) scenes," Proc. Soc. Photo-Opt. Instrum. Eng. **178**, 13 (1979).

10. J.J. Otazo, E.W. Tung, and R.R. Parenti, "Digital filters for infrared target acquisition sensors," Proc. Soc. Photo-Opt. Instrum. Eng. **238**, 78 (1980).

11. I.W. Kay, "Evaluation of infrared target discrimination algorithms," IDA Paper P-1714 (Institute for Defense Analyses, Alexandria, VA, 1983).

12. N. Ben-Yosef, K. Wilner, S. Simhony, and G. Feigin, "Measurement and analysis of 2-D infrared natural background," Appl. Opt. **24**, 2109 (1985).

13. Y. Itakura, S. Tsutsumi, and T. Takagi, "Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region," Infrared Phys. **14**, 17 (1974).

14. W.N. Peters, "One- and two-dimensional matched filter for scanning electro-optic systems," J. Opt. Soc. Am. A **3**, 347 (1986).

15. E.H. Takken, D. Friedman, A.F. Milton, and R. Nitzberg, "Least-mean-square spatial filter for IR sensors," Appl. Opt. **18**, 4210 (1979).

16. B. Widrow, J.M. McCool, M.G. Larimore, and C.R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," Proc. IEEE **64**, 1151 (1976).

17. B. Widrow, "Adaptive Filters," pp. 573-581 in *Aspects of Network and System Theory* (Holt, Rinehart and Winston, New York, 1971), edited by R.E. Kalman and N. DeClaris.

18. B. Widrow and S.D. Stearns, *Adaptive Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1985), Ch. 6.

19. H. Stark and F.B. Tuteur, *Modern Electrical Communications: Theory and Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1979), Sec. 11.3.

20. R.J. Schwarz and B. Friedland, *Linear Systems* (McGraw-Hill, New York, 1965).

21. C.E. Cook and M. Bernfeld, *Radar Signals; Introduction to Theory and Application* (Academic Press, New York, 1967).

22. H.J. Blinchikoff and A.I. Zverev, *Filtering in the Time and Frequency Domains* (Wiley, New York, 1976).

23. M.S. Longmire, A.F. Milton, and E.H. Takken, "Simulation of mid-infrared clutter rejection. 1: One-dimensional LMS spatial filter and adaptive threshold algorithms," Appl. Opt. **21**, 3819 (1982).

24. M.S. Longmire, F.D. Bryant, J.D. Wilkey, and A.F. Milton, "Simulation of mid-infrared clutter rejection. 2: Threshold-sensor-size effects with LMS-filtered noise," Appl. Opt. **24**, 579 (1985).

25. P.S. Maybeck, *Stochastic Models, Estimation and Control* (Academic Press, New York, 1979), Vol. 1, Secs. 4.3, 5.12.

26. R.W. Harris and T.J. Ledwidge, *Introduction to Noise Analysis* (Pion, London, 1974), Secs. 8.2, 8.3.

27. J.S. Bendat and A.G. Piersol, *Random Data: Analysis and Measurement Procedures* (Wiley, New York, 1971).

28. R.N. Bracewell, *The Fourier Transform and Its Applications* (McGraw-Hill, New York, 1978), 2nd ed.

29. A.P. Schaum, Naval Research Laboratory, "Signal processing methods for sensitivity enhancement in a staggered-column focal plane array," private communication (1986).

30. S. Tutumi, "Spatial filter used in scanning optical detection system," Electronics and Communications in Japan **49**, no. 6, 13 (1966).

31. J.S. Bendat, *Principles and Applications of Random Noise Theory* (Wiley, New York, 1958).

32. J.H. Eberly and K. Wódkiewicz, "The time-dependent physical spectrum of light," J. Opt. Soc. Am. **67**, 1252 (1977).

33. W.K. Pratt, "Vector space formulation of two-dimensional signal processing operations," J. Comput. Graphics Image Proc. **4**, 1 (1975).

34. R.M. Mersereau and D.E. Dudgeon, "The representation of two-dimensional sequences as one-dimensional sequences," IEEE Trans. Acoustics, Speech, and Signal Processing **ASSP-22**, 320 (1974).

35. J.B. Scarborough, *Numerical Mathematical Analysis* (Johns Hopkins Press, Baltimore, MD, 1966), 6th ed.

36. E. Whittaker and G. Robinson, *The Calculus of Observations* (Dover, New York, 1967), 4th ed.

37. M.V. Wilkes, *A Short Introduction to Numerical Analysis* (Cambridge U. Press, Cambridge, 1966), pp. 48-51, Ch. 6.

38. C.F. Gerald, *Applied Numerical Analysis* (Addison-Wesley, Reading, MA, 1978), 2nd ed., Secs. 4.1-4.4.

39. F.B. Hildebrand, *Introduction to Numerical Analysis* (McGraw-Hill, New York, 1956), Secs. 3.3, 5.3.

40. D.R. Hartree, *Numerical Analysis* (Clarendon Press, Oxford, 1958), 2nd ed., Secs. 4.5-4.7, 6.7, 10.2, 10.3.

41. J.F. Steffensen, *Interpolation* (Chelsea Pub. Co., New York, 1950), 2nd ed., Sec. 7.

42. D.M. Young and R.T. Gregory, *A Survey of Numerical Mathematics* (Addison-Wesley, Reading, MA, 1972), Secs. 7.1-7.3.

43. W.H. Beyer (ed.), *Standard Mathematical Tables* (CRC Press, Boca Raton, FL, 1984), 27th ed., Ch. XI.

44. R.V. Southwell, *Relaxation Methods in Theoretical Physics* (Clarendon Press, Oxford, London, 1946), vol. 1, pp. 229-237.

45. For another method of indexing and deriving 2-D operator arrays, see W.G. Bickley, "Finite difference formulae for the square lattice," Quart. J. Mech. and Applied Math. 1, 35 (1948).

END

FILMED

MARCH, 1988

DTIC